The First

caterpillars,
'karyotes, &
consciousness

Arthur S. Reber

Primele minți

omizi, „cariote” și conștiință

de

Arthur S. Reber

Prefață
Unde a început călătoria: O notă personală
O scurtă prezentare generală
O paranteză asupra terminologiei
O paranteză despre natura argumentului filozofic
Distincția între biologia evoluționistă și
psihologia evoluționistă
O notă privind citatele
Rezumat
Recunoștințe

1
Pot roboții să fie conștienți?
Funcționalismul sau posibilitatea unei conștiințe
„independente de hardware”
Ce știu eu/voi despre minți - a ta și/sau a mea?
Există ceva asemănător cu a fi robot?
John Searle's Chinese Room
The Anthropomorphic „Tug”
Rezumat

2
De ce nu a fost propus înainte un model de tip bază celulară a conștiinței?
Omida mea propedeutică
Trei motive de ezitare
În apărarea antropomorfismului
O scurtă paranteză despre durerea peștilor - o lecție de obiect
Cei care sunt. Ar putea fi, sau poate ar trebui să fie confortabili cu celulele simțitoare
Rezumat

3
Viziuni biocentrice asupra originii conștiinței: Căutarea soluției la o problemă dificilă
Teorii generale ale originii conștiinței
Căutarea phvlogenetică a primei apariții a conștiinței
Câteva reflecții asupra acestor eforturi
Se împiedică de o eroare de categorie
Dilema emergentiștilor
Anticiparea criticilor: Unele argumente profilactice
Rezumat

4
Când devenim serioși: bazele celulare ale conștiinței
Puncte comune cu alte domenii de cercetare
O scurtă paranteză despre un moment de nedumerire personală
O scurtă paranteză despre care specie unicelulară vorbim
Cadrul bazei celulare a conștiinței
O privire mai atentă asupra vieții unicelulare
Rezumat

5
Coda: Anticiparea criticilor (încă o dată) şi apărarea împotriva lor
Aspecte ale poziţiei bazei celulare a conştiinţei
Cu privire la structura ierarhică a evoluţiei
Reducţionismul
Evitarea misterianismului
O scurtă remarcă asupra dualismului
Probleme empirice în căutare de soluţii
Cum rămâne cu plantele?
Consideraţii etice
O scurtă introducere despre altruism şi cooperare
Câştiguri secundare
Rezumat

Anexa A:
Un exerciţiu de lexicografie: Definirea (?) conştiinţei

Apendicele B:
Comentariu privind Declaraţia de la Cambridge privind conştiinţa

Prefață

Unde a început călătoria: O notă personală

Fiecare călătorie intelectuală are un început, un prim pas, real sau abstract. Poate fi o intuiție bruscă, un indiciu lăsat să atârne dintr-o conversație auzită din întâmplare, o intuiție neobservată care clocește chiar sub pragul conștiinței sau, în cazul meu, un deceniu de meditații vagi care s-au cristalizat în primăvara anului 1993, în timpul unei discuții improvizate și în mare parte unidirecționale cu o omidă mică și verde care îmi distrugea plantele de busuioc.

Figura P.1 oferă o imagine a prietenului meu poliped, iar figura P.2 a ceea ce el/ea[1] avea să devină în curând.

Figura P. 1. Prietenul meu poliped (tehnic, Trichoplusia ni) în stadiul de larvă. Cu permisiune. Credit foto, Phil Bendle, curator, T.E.R.R. A.I.N: Taranaki Educational Resource Research Analysis and Information Network, Noua Zeelandă.

Figura P.2. T. ni în faza sa postcocoon ca lucrător al verzei. Cu permisiune. Credit foto, Ian Maton, BugGuide.net, Departamentul de Entomologie, Universitatea de Stat din Iowa.

1 În timp ce cercetam diverse domenii obscure ale academiei, am descoperit că este aproape imposibil să determinăm sexul majorității omizilor fără disecție, ceea ce nu era o opțiune în acest caz.

Interacţiunea a avut un impact puternic asupra mea. Un sfert de secol mai târziu, încă mai pot vedea acea molie incipientă, corpul său articulat clătinându-se deasupra unei frunze de busuioc. Am făcut tot posibilul să o imortalizez cu titlul acestui volum. Cred că vă va plăcea de ea (sau de el).

Aceasta este o carte scurtă. Are un mesaj direct. Voi prezenta o teorie despre originile conştiinţei şi ale minţii, pe care am numit-o **Baza celulară a conştiinţei** (CBC). Aceasta se bazează pe câteva ipoteze foarte simple, şi anume că:

> *(i) originile minţii şi conştiinţei se vor găsi în cele mai simple organisme unicelulare şi,*
> *(ii) Conştiinţa, subiectivitatea, experienţa fenomenală sau, dacă preferaţi, sensibilitatea este o caracteristică inerentă a formei (formelor) organice vii.*

Pentru a anticipa, da, voi susţine că speciile unicelulare precum *amoebae* au minţi, deşi sunt foarte mici şi nu fac prea multe; că protozoarele percep lumea din jurul lor şi gândesc, deşi gândurile lor sunt limitate şi nu foarte interesante; că bacteriile comunică între ele, deşi mesajele sunt simple şi unitare prin natura lor; şi că eucariotele sesile precum *Stentor roeseli*[2] nu numai că învaţă, dar au memorii unicelulare minuscule şi iau decizii tactice. Şi că, da, toate posedă o conştiinţă ontologic valabilă.

Deşi unii pot considera aceste propuneri ca fiind radicale, eu nu. Obişnuiam să o fac. Nu o mai fac. După ce terminăm, sper că nici dumneavoastră nu o veţi mai face. Aşa cum voi susţine, aceste principii sunt evidente din punct de vedere intuitiv, seducătoare din punct de vedere intelectual şi au sens dintr-o perspectivă biologică evolutivă. Sunt, de asemenea, noţiuni cu care destul de mulţi biologi, neuroştiinţifici şi filozofi s-au jucat din când în când de ceva timp, deşi majoritatea au refuzat în cele din urmă, aşa cum se spune în jocuri precum *backgammon*-ul, opţiunea. Eu am fost, în zilele mele *precaterpillar,* un membru semnat al acestui grup de declinatori. În cartea mea din 1993, *Implicit Learning and Tacit Knowledge*, am scris,

Pare destul de clar că până şi cele mai primitive organisme pot fi „creditate" cu unele forme rudimentare de receptivitate la mediu; chiar şi organismele unicelulare reacţionează la stimuli nocivi şi fac locomoţii „deliberative" pentru a se îndepărta de aceştia. Cu toate acestea, nu este clar dacă dorim să recunoaştem că un <u>Paramecium</u> are conştiinţă atunci când se retrage (observaţi, nu am spus „evită") de la un stimul nociv. Eu, unul, sunt mulţumit să las astfel de organisme într-un fel de stare carteziană mecanicistă; repertoriile lor comportamentale par să fie destul de reflexe, iar Canonul lui Lloyd Morgan ar trebui să fie ghidul nostru pentru înţelegere.

Reticenţa mea nu se datora faptului că noţiunea unui *Paramecium* cu adevărat conştient nu avea o bază ştiinţifică, ci îngrijorării de a părea că îmbrăţişez idei nebuneşti. După cum mi-a povestit respectatul microbiolog britanic Dennis Bray, prima dată când a evocat într-un forum public posibilitatea unei *amoebe* conştiente, sugestia a fost întâmpinată cu un val de batjocură şi respingere atât de intens, încât atunci când a scris cartea sa fundamentală din 2009, *Wetware*, el, ca şi mine în 1993, a respins în mod explicit posibilitatea conştiinţei la speciile unicelulare.

La doi ani după întâlnirea mea cu ceea ce descoperisem până atunci că era un specimen din specia *Trichoplusia ni,* am fost invitat să iau cuvântul în cadrul Asociaţiei Britanice de Psihologie. Nu se întâmplă prea des să primeşti o astfel de invitaţie şi am decis că ar fi momentul potrivit să o fac publică. Mă simţeam mai sigur pe gândirea mea şi eram pregătit să rezist la ceea ce am presupus că vor fi hărmălaia britanică clasică şi atacurile inteligent meşteşugite la adresa sănătăţii mele mintale, împletite cu un amuzament vag deghizat cu privire la o posibilă demenţă precoce. Cu multă nerăbdare, am prezentat o primă versiune a ceea ce mai târziu s-a transformat în modelul **CBC**.

---

2 Un eucariot acvatic unicelular a cărui gamă remarcabilă de comportamente şi funcţii a fost catalogată pentru prima dată de microbiologul Herbert Jennings în cartea sa clasică din 1906, *The Behavior of the Lower Organisms*. Îl vom analiza mai îndeaproape mai târziu.

Ceea ce s-a întâmplat a fost una dintre cele mai ciudate experienţe din viaţa mea în academie. Mi-am ţinut discursul în faţa celor adunaţi. A fost primită cu aplauze politicoase, apoi toată lumea s-a ridicat şi a plecat. Nici măcar o întrebare. Niciun efort de tâlhărie intelectuală. Şi, bineînţeles, niciun semn de aprobare sau de recunoaştere a curajului meu. „Ce s-a întâmplat?" Am întrebat-o pe preşedinta sesiunii. „Ah", a zâmbit ea cu melancolie, "*cred că ar fi trebuit să vă spun. Am avut o mică dispută anul trecut. Un prezentator a fost supus unui atac destul de neplăcut şi necruţător care l-a lăsat destul de zdruncinat. Am decis că, anul acesta, dacă un subiect delicat se află pe ordinea de zi, iar dvs. era, nu vor exista întrebări sau răspunsuri*".[3]

Aşadar, am avut parte de încă unul dintre acele momente de viaţă delicios de ciudate. Eram aici, dornic de luptă, căutând să am un schimb de replici bun cu acea mănuşă renumită, o adunare a academicienilor britanici. Am rămas cu acea senzaţie vagă de gol pe care o ai când ai inhalat un pahar de sifon acidulat. Tot ce iei în plămâni este dioxid de carbon şi te lasă cu senzaţia că lipseşte ceva.

Aşa că am decis să încerc calea revistei. Am trimis o versiune a discursului la revista interdisciplinară *Philosophical Psychology.* După obişnuitul du-te-vino cu recenzorii, „*Caterpillars and Consciousness*" a fost publicată în 1997. Şi asta a fost cam ultima oară când cineva a văzut-o. O verificare a indexului de citări a arătat că, în anii următori, a fost menţionată de exact trei autori... şi unul dintre ei eram eu.

Dar ideile nu mor singure. Fie sunt carotate de colegii critici, fie sunt scoase din scenă de autorii lor. Eu nu am fost supus celei dintâi situaţii pentru că nimeni nu mi-a acordat atenţie şi nu m-am retras de pe scenă, ci pur şi simplu am mutat ideea de bază în acel loc clişeu, pe locul din spate. Am cultivat propunerea iniţială doar pentru că am crezut că am dreptate (şi da, văd aroganţa din asta). Dar lipsa totală de recunoaştere, faptul că, în esenţă, nimeni nu a recunoscut propunerea de bază după ce a fost publicată într-o revistă interdisciplinară respectată, m-a împiedicat să revin asupra problemei - până de curând.

În 2016, Stevan Hamad a început să editeze o nouă revistă, *Animal Sentience*. L-am contactat pe Hamad pentru a-l întreba dacă crede că o actualizare a cărţii „Caterpillars" ar fi un lucru rezonabil de făcut. S-a dovedit că el era unul dintre foarte puţinii neuroştiinţifici care ştiau despre lucrarea originală. Este foarte posibil ca acesta să fi fost prezent la reuniunea BPA din 1995, având în vedere că la acea vreme făcea parte din corpul profesoral al Universităţii din Southampton (acum este Emerit acolo şi îşi petrece timpul la Universite du Quebec a Montreal[4]). El a spus „sigur", avertizându-mă că nu numai că va fi supus obişnuitei evaluări *inter pares*, dar va deveni un articol ţintă. La fel ca în cazul revistei *Behavior and Brain Sciences*, pe care Harnad a înfiinţat-o în 1976, *Animal Sentience* folosea un format de comentariu deschis, în care orice lucrare publicată poate fi criticată de oricine (deşi, desigur, aceste comentarii sunt la rândul lor supuse evaluării colegiale).

Caterpillars, Consciousness and the Origins of Mind a apărut în 2016 şi a fost criticată rapid de o jumătate de duzină de academicieni din filosofie, biologie şi psihologie - exact ceea ce am sperat cu două decenii mai devreme. Am urmat cu un răspuns (Reber, 2016b) şi am fost întâmpinat cu un alt set de critici atente foarte binevenite care, desigur, au primit un alt răspuns (Reber. 2017). Toată această activitate a condus aici, cu Oxford University Press. Deoarece multe dintre ideile de bază care vor fi dezvoltate în paginile care urmează sunt în consonanţă cu materialul din aceste publicaţii anterioare, voi împrumuta din când în când din ele, rezultând un oarecare „autoplagiat". Este genul de lucru care tinde să se strecoare printre scriitorii care se îndrăgostesc de cuvintele lor. Mă declar vinovat pentru acest mic păcat. Aceasta ne va permite, de exemplu, să ocolim o serie de probleme aproape irezolvabile care au însoţit alte abordări ale studiului conştiinţei şi, în mod critic, ne va permite să restructurăm multe dintre întrebările esenţiale care au apărut în domenii atât de

3 O colegă care a fost prezentă îşi aminteşte acest lucru diferit. Ea crede că moderatorul a spus că în anul precedent nimeni nu a pus întrebări, iar vorbitorul a fost captivat de o tăcere atât de dureroasă, încât au decis să prevină orice jenă viitoare, pur şi simplu având o politică de interdicţie a întrebărilor. În orice caz, rezultatul a fost cel care a fost.

4 Sau, aşa cum este adesea menţionată printre anglofoni, Universitatea din Montreal, deşi, de fapt, aceasta este o instituţie separată.

diverse precum filosofia minții, biofilosofia, inteligența artificială (AI), robotica, psihologia cognitivă, neuroștiințele și biologia evolutivă. Acestea sunt lucruri amețitoare și reprezintă o provocare de a le prezenta în moduri care să atragă acea cohortă favorizată, a profanilor inteligenți, menținând în același timp poziții coerente din punct de vedere științific și filosofic cu care savanții și cercetătorii se vor simți confortabil. Sincer, habar nu am dacă am reușit acest lucru - dar bănuiesc că voi afla în curând.

O prezentare rapidă

Acest mic volum este un efort de a răspunde la câteva dintre acele întrebări care ne frământă pe toți dintotdeauna. De ce mintea umană? De ce avem un sentiment atât de puternic al sinelui? De ce suntem conștienți, nu doar de lucrurile din lume, ci și de faptul că suntem conștienți de acele lucruri? Care sunt originile acestor diverse abilități cognitive și afective care sunt adesea grupate sub umbrela conștiinței? Și cea mai importantă: Cum poate creierul uman, un obiect biofizic, să creeze gânduri și emoții, conștiință, subiectivitate, experiență? Cum poate materialul să creeze imaterialul?

Prin abordarea acestor probleme, mă alătur unei liste foarte lungi de gânditori care au scris cărți, articole, tratate, disertații, precum și predici, simfonii și poezii, toate având același scop. Dar există o diferență, una semnificativă. Aceste alte încercări s-au concentrat pe rezolvarea „problemei conștiinței”. Eu nu voi încerca asta. În schimb, voi reformula întrebările, ceea ce ne va permite să rezolvăm problema. După cum vom vedea, practic toate celelalte eforturi de a răspunde la întrebările privind originile minții și conștiinței încep cu creierul uman, mintea umană și cogniția umană. Modelul **CBC** recomandă o mișcare de inginerie inversă - demontați totul, începeți cu componentele originale, în acest caz, cele mai simple organisme și urcați pe scara evoluției.

În prima parte a cărții, vom analiza alte două abordări ample care au fost adoptate pentru aceste probleme. Una provine din domeniul inteligenței artificiale, al roboticii și din unele domenii ale filosofiei minții. Aceasta ia în considerare posibilitatea ca mintea și conștiința să nu fie neapărat caracteristici ale speciilor organice pe bază de carbon, ci ar putea fi captate de artefacte, entități computaționale sofisticate. Vom examina și vom critica această poziție. Deși această abordare suscită încă interesul scriitorilor de *science-fiction* și al futuriștilor, în ultimele două sau trei decenii și-a pierdut o mare parte, dar nu toată, cu siguranță nu toată, susținerea în cadrul comunității științifice.[5] Motivele acestei schimbări de orientare și ale dezinteresului crescând în rândul oamenilor de știință și al filosofilor minții sunt interesante și le vom analiza îndeaproape.

Cealaltă abordare este cea care a dominat gândirea și cercetarea în filosofia minții și în neuroștiințe în ultimele câteva decenii. Aici, conștiința este vizată prin formularea unor întrebări cu un caracter neurocognitiv clar, cum ar fi: „*Cum face creierul mintea?*” Sau: „*Cum pot operațiile, funcțiile și procesele neuronale să cauzeze conștiința?*” Sau: „*Cum poate fi experiența subiectivă evanescentă, nonfizică, rezultatul unei «simple» acțiuni biochimice?*” Sau „*Cum au evoluat creierele astfel încât la un moment dat să apară conștiința de sine?*” Sau... Ei bine, ați înțeles ideea - aici se caută identificarea căilor neuronale, a centrelor și a structurilor din creier care funcționează astfel încât să dea naștere stărilor fenomenale.

---

5 Într-adevăr, în edițiile recente ale revistei Science, cercetătorul în științe cognitive Olivia Carter și colegii săi au avut un schimb de opinii interesant cu Stanislaw Dehaene și coautorii săi cu privire la posibilitatea inteligenței artificiale conștiente, ambele părți concluzionând că obiectivul a ceea ce se numește „inteligență artificială puternică” este încă posibil, că într-o zi ar putea fi construite artefacte conștiente. Nicolas Spatola și Karolina Urbanska au intrat, de asemenea, în luptă, susținând că roboții conștienți ar trebui să aibă aceleași drepturi ca oamenii. Deși nu sunt de acord cu premisa de bază, argumentele prezentate mi s-au părut interesante și le vom aborda în detaliu în capitolele 1 și 5.

Luate în ansamblu, aceste întrebări reprezintă ceea ce filosoful David Chalmers a numit în mod celebru *„problema grea a conștiinței”* și a devenit una dintre cele mai tentante enigme nerezolvate de pe scena științifică. În 2005, *Science*, una dintre cele mai respectate reviste interdisciplinare din lume, a clasificat întrebarea *„Care este baza biologică a conștiinței?”* drept a doua cea mai importantă dintre întrebările științifice rămase fără răspuns. Aceasta a fost clasată doar după întrebarea *„Din ce este făcut universul?”* și înaintea întrebărilor *„De ce oamenii au atât de puține gene?”* și *„Pot fi unificate legile fizicii?”* Vom examina și vom critica această abordare. Critica se va concentra pe unele probleme inerente pe care mulți cercetători din domeniu par să nu le fi recunoscut. De exemplu, pe măsură ce cercetările au progresat, s-a descoperit că un număr tot mai mare de specii au funcții cognitive mai sofisticate decât bănuiau primii etologi și biologi evolutivi, iar întrebările cu privire la unicitatea conștiinței umane sunt foarte actuale.

În 2012, la *Conferința comemorativă Francis Crick* privind conștiința la animalele umane și non-umane, desfășurată la Cambridge, Regatul Unit, un grup de neuroștiințifici proeminenți au recunoscut oficial importanța acestei probleme și au semnat *Declarația de la Cambridge* privind conștiința:

Absența unui neocortex nu pare să împiedice un organism să experimenteze stări afective. Dovezile convergente indică faptul că animalele neumane au substraturile neuroanatomice, neurochimice și neurofiziologice ale stărilor conștiente, precum și capacitatea de a manifesta comportamente intenționate. În consecință, toate dovezile indică faptul că oamenii nu sunt singurii care posedă substraturile neurologice care generează conștiința. Animalele non-umane, inclusiv toate mamiferele și păsările și multe alte creaturi, inclusiv caracatițele, posedă, de asemenea, aceste substraturi neurologice.

Declarația a fost elaborată de neurobiologul computațional Philip Low și editată de un grup de cercetători respectați, printre care regretatul Jaak Panksepp, specialistul în cetacee Diana Reiss și neurocercetătorii David Edelman, Bruno Van Swinderen și Christof Koch. (Vom întâlni mai mulți membri ai acestui grup mai târziu, deoarece munca lor este relevantă pentru teza mea). Acesta a fost urmat, de asemenea, de o serie de comentarii privind problema generală a conștiinței, inclusiv ce specii pot fi confortabil acordate să o manifeste și cum ne încadrăm noi, ca oameni, în panoplia completă a vieții de pe această planetă. A fost un efort nobil, dar a suferit de o dificultate care, după cum vom vedea, se inserează în aproape toate conversațiile, discuțiile și dezbaterile pe această temă. Scriitorii au început cu varietatea conștiinței prezente în noi, *Homo sapiens sapiens*, și apoi și-au întors privirea înapoi spre alte specii, căutând dovezi ale tipului nostru de sensibilitate. Atunci când au apărut dovezi în acest sens, acelei specii i-au fost acordate stări mentale ontologice. În cazul în care acestea nu au apărut, nu li s-au acordat. Acest model de gândire poate fi văzut în mod clar în textul integral al acestor comentarii, care este reprodus în apendicele B.

Eforturile de a utiliza această strategie bioevolutivă, și vom analiza mai mult decât câteva dintre ele, au produs o gamă destul de largă de cazuri în care unul sau altul dintre cercetători crede că a identificat punctul de-a lungul continuumului evolutiv în care, sau la care specie, a apărut acest lucru al conștiinței. Pisicile și câinii o au? Majoritatea cred că da. Dar unele dintre acele specii aviare inteligente, cum ar fi corbii, ciorile și alte corvide? Din nou, există un acord considerabil. Este posibil ca unele specii acvatice, cum ar fi caracatițele și alte cefalopode, care s-au detașat pe propria ramură evolutivă și au dezvoltat tipuri foarte diferite de corpuri și sisteme nervoase, să aibă o viață mentală autentică? Aici, există mai puțin acord. Cum rămâne cu insectele? Moluștele? Bacteriile?

Pe lângă faptul că a creat multe dispute între cercetători cu privire la modul în care aceste date sunt interpretate și la ceea ce „contează”, cu adevărat și fără ambiguitate, ca dovadă a conștiinței, există o problemă și mai mare - pe care o voi numi ***Dilema Emergentiștilor***. Nu este

vorba pur și simplu de acordarea unei specii (sau clade sau phylum) o conștiință reală din punct de vedere ontologic; trebuie elaborat un argument coerent care să arate de ce, atunci când mecanismele evolutive au produs în cele din urmă o specie cu sensibilitate, conștiință, acest lucru asemănător minții a izbucnit în existență în timp ce, cu un moment cosmic mai devreme, toate speciile de complexitate mai redusă erau simple entități mecanice care se comportau fără nici măcar o licărire de conștiință.[6] Vom analiza această problemă în detaliu, deoarece ea dezvăluie ceea ce, în opinia mea, este un defect fatal în toate aceste abordări.

Poziția **CBC**, desigur, are propria sa dilemă emergentistă și nu poate evita problema. Ei bine, trebuie să o abordăm cât de sincer putem. Voi susține că avantajul de a adopta perspectiva **CBC** este că miracolul emergentist pe care îl necesită este unul mult mai ușor de rezolvat decât cel cu care se confruntă toate celelalte abordări. Dintr-o perspectivă biologică evolutivă, descoperirea modului în care sensibilitatea a început chiar de la început, cum a apărut odată cu prima pâlpâire de viață, va fi o problemă mult mai ușor de rezolvat decât încercarea de a determina mai întâi care specii sunt cu adevărat sensibile și apoi lupta pentru a descifra neurobiologia de bază care distinge aceste specii de cele care le-au precedat sau s-au rătăcit pe alte ramuri evolutive.

Cu toate acestea, critica mea față de o mare parte a cercetării contemporane în aceste domenii nu va fi prezentată ca un atac la adresa numeroaselor programe științifice minunate care explorează elementele neurocognitive care stau la baza conștiinței umane și a funcționării mentale umane. Sunt un fan al științei creierului și mă simt bine pe dinăuntru de fiecare dată când se face o nouă descoperire care arată cum o anumită structură cerebrală modulează o anumită capacitate cognitivă de nivel înalt. De fapt, ne vom uita la unele dintre acestea, dar cu înțelegerea clară că identificarea mecanismelor neuronale pentru, să zicem, capacitatea de a reprezenta sinele nu este răspunsul la întrebarea: „*Cum face creierul să determine mintea să capteze sensul fenomenal al sinelui?*” Răspunsul la această întrebare, susțin eu, va fi găsit doar atunci când vom începe prin a descifra modul în care speciile unicelulare se pot identifica și pot deosebi, ca să spun așa, pe mine de tine. Și, după cum vom vedea, ele o fac.

A doua parte a proiectului va consta în dezvoltarea mai detaliată a teoriei **CBC**. Vom arăta cum se încadrează în abordările standard ale biologiei evolutive și cum se pliază perfect pe o schelă teoretică pe care am folosit-o anterior pentru a dezvolta o teorie a inconștientului cognitiv bazată pe evoluție (pentru detalii, a se vedea Reber, 1993). Mă voi ocupa, de asemenea, de „problema dificilă” personală - arătând de ce are sens să considerăm că cele mai primitive forme de viață au experiențe interne, subiective, legitime, ceva ce putem numi minte fără a părea nebunesc.

Ne vom uita la biologia detaliată a speciilor unicelulare și vom explora sistemele lor surprinzător de sofisticate de detectare, reprezentare, învățare, memorare și comunicare cu lumea înconjurătoare. Vom analiza modul în care acestea stabilesc valențele obiectelor pe care le întâlnesc pentru a determina ce este comestibil și nutritiv și ce este nociv și potențial dăunător pentru țesuturi. Vom explora capacitatea lor de a învăța din experiență și vom dezvălui dovezi convingătoare ale unor capacități memoriale minime, dar reale. Vom analiza, de asemenea, dovezile privind comunicarea atât în cadrul speciilor unicelulare, cât și între acestea, ceea ce permite comportamente adaptative, cum ar fi formarea de colective care prezintă comportamente sociale precum roiurile și ajustările dimensiunii populației care sunt sensibile la nivelurile de nutrienți. Vom explora, de asemenea, dovezi convingătoare ale unui altruism rudimentar, în care coloniile de bacterii se pun în pericol pentru a-și ajuta vecinii aflați în nevoie. Aceste funcții sunt de obicei identificate ca aspecte esențiale, componente ale sensibilității. Toate acestea se regăsesc în cele mai elementare forme de viață. Ele nu vorbesc doar despre acțiuni mecanice sau comportamente lipsite de afect, ci despre o experiență fenomenală autentică, despre subiectivitate și minte. Voi susține că există ceva ce înseamnă să fii un *Paramecium* sau o *amibă* sau o *bacterie*.

6 Înțelegând că „izbucnirea în existență” este, în termeni biologici evolutivi, un proces care are loc pe perioade lungi de timp.

Pe parcurs, vom aborda, de asemenea, o serie de aspecte conexe din filosofia științei și a minții. Adoptarea poziției solicitate de **CBC** se dovedește a avea unele implicații intrigante, în special pentru filosofia minții. Vom examina aici o varietate de *„**-isme**”*, inclusiv emergentismul, reducționismul, fizicalismul, dualismul, misterianismul și etica, și vom pune întrebări precum *„Pot fi plantele sensibile?*” De asemenea, vom aborda, într-un mod speculativ, motivul pentru care există acest decalaj uriaș în ceea ce privește capacitatea mentală între *Homo sapiens* și toate celelalte specii care, după cum am citit literatura, este cel puțin parțial responsabil pentru abordarea centrată pe om care a dominat atât filosofia, cât și științele cognitive.

Filosoful influent Daniel Dennett a susținut că distanța conceptuală dintre conștiința umană și funcțiile mentale și cele ale rudelor noastre cele mai apropiate este la fel de mare (sau poate mai mare) decât cea dintre limbajul uman și cântecul păsărilor.

Bănuiesc că această prăpastie funcțională poate fi pusă pe seama faptului că toate formele intermediare de hominizi care ar fi putut prezenta un limbaj rudimentar, utilizarea uneltelor, autoreferința și alte aspecte ale formelor noastre „speciale” de sensibilitate au dispărut. Având în vedere că speciile care ar fi putut prezenta capacități „apropiate de noi” sunt toate bine îngropate, devine greu să nu te implici într-o criză de narcisism al speciei.

O introducere în terminologie

Noțiunea de conștiință umană, cu autocunoașterea sa curajoasă și proprietățile sale reflexive, este un subiect seducător, atrăgător și irezistibil. Este, de asemenea, una dintre cele mai vechi probleme cu care se confruntă cercetătorii. A fost abordată de sute, mii, zeci de mii de poeți, eseiști, romancieri, artiști, gânditori și oameni de știință, de la macrobiologi la teoreticieni cuantici; de microbiologi, psihanaliști, neuroștiințifici cognitivi și cercetători în domeniul inteligenței artificiale; de biologi evolutivi, geneticieni, fantezişti futuriști și scriitori de *science fiction*; de teologi și, bineînțeles, de filosofi ai minții, care se comportă adesea ca și cum ar fi inventat (și, prin urmare, dețin) subiectul.

Impactul acestei abundențe literare și științifice a fost exact ceea ce te-ai aștepta. Există o literatură uriașă pe aceste teme și, deoarece interesul a venit din partea atâtor persoane cu o varietate atât de mare de medii, credințe și obiective, termeni precum conștiință, minte, sensibilitate, subiectivitate, fenomenologie, stare mentală, experiență, *qualia* (și mulți alții) sunt răspândiți în mod nebunesc. Ocazional, puriștii lexicografici[7] încearcă să țină în frâu discursul, insistând asupra unor definiții clare ale termenilor sau impunând limite conotative stricte pentru utilizare. Unii refuză să utilizeze conștiința pentru orice altă specie în afară de a noastră. Alții se străduiesc să identifice un prag, un punct în complexitatea biologică în care se pot simți siguri că o specie este dotată cu o minte. Una dintre cele mai citate lucrări de filosofie este *„Cum e să fii liliac?*” a lui Thomas Nagel, care a încercat cu curaj să abordeze această problemă. Alții se plâng (și da, scrierile lor au adesea senzația că autorul se plânge în timp ce compune) că, până când nu clarificăm ce înseamnă acești termeni, vom rămâne blocați, împotmoliți într-o confuzie lingvistică autoprovocată.

Mă retrag din această problemă - ca lexicograf practicant, îmi revendic dreptul de a face acest lucru. Sunt de părere că, pentru scopurile mele aici, un astfel de efort de definire, oricât de bine intenționat ar fi, nu este doar fără speranță, nu este nici măcar plin de speranță. Nu sunt singurul în această situație. După cum a remarcat neurologul cognitiv Christof Koch în meditația sa adesea personală pe această temă (*Consciousness: Confessions of a Romantic Reductionist*), de multe ori știința progresează mai eficient atunci când termenii sunt utilizați într-o manieră *ad hoc*, pentru că ceea ce poate părea a fi claritatea necesară nu face decât să restrângă viziunea și să limiteze tipurile de întrebări care ar putea fi puse. Puțină neclaritate poate fi utilă. Pentru cei cărora le plac aceste mici rătăciri lingvistice, aruncați o privire la *apendicele A*. Este o trecere în revistă a

7 Numărați-mă printre ei. A se vedea Reber (1985). Reber & Reber (2001). și Reber. Allen & Reber (2009).

diverşilor fluturaşi conotativi în care s-au blocat lexicografii încercând să definească termeni precum conştiinţă şi minte şi, *atenţie sporită*, eu sunt unul dintre ei.

Pentru restul acestei cărţi, voi lăsa deoparte afilierea mea cu colegii mei care muncesc ca ceea ce lexicograful nostru Samuel Johnson ne numea: „truditori utili”. Mi-am învăţat lecţia. Am de gând să mă joc rapid şi liber cu cuvintele. Termeni precum conştient şi conştiinţă, minte şi mental, conştient şi conştiinţă, sensibil şi sensibilitate, subiectiv şi subiectivitate, mental şi stare mentală, percepţie şi perceptual, experienţă şi experienţial, fenomenal şi fenomenologic (şi alţii, mulţi alţii) vor fi utilizaţi în moduri vag sinonime. Acestea îşi vor exprima semnificaţiile în ceea ce se numeşte adesea o manieră de „psihologie populară” - şi anume, modul în care nespecialiştii vorbesc despre aceste subiecte, independent de impunerile poliţiei lingvistice a academiei. Dacă există un termen de acoperire aici, acesta este probabil *sentimentul* care, după cum s-a menţionat mai sus, se regăseşte şi în numele noii reviste, *Animal Sentience*, în care probleme precum cele abordate în această carte pot fi întâlnite în mod regulat.

Voi adopta o poziţie similară în ceea ce priveşte termenii filosofici precum ontologie, epistemologie şi existenţial, precum şi derivatele acestora. Îi voi folosi destul de liber, încercând în acelaşi timp să rămân în limitele lexicografice ale specialiştilor. Ontologia se referă în general la o abordare a metafizicii care explorează fiinţa sau devenirea, existenţa, realitatea - şi cam aşa o voi folosi eu. O frază precum „sensibilitatea unei omizi este stabilită ontologic” înseamnă că dovezile sunt suficient de puternice şi solide pentru a concluziona că aceasta are experienţe subiective interne şi că acestea sunt reale. O ontologie este un cadru mai larg în care sunt clasificate şi organizate categoriile de obiecte sau alte domenii de cunoaştere. Nu ne va folosi prea mult acest sens aici, în afară de presupunerea că toate formele de sensibilitate alcătuiesc o ontologie fenomenologică. Epistemologia este o ramură a filosofiei care se ocupă cu cunoaşterea, reprezentarea acesteia (adică forma în care este păstrată într-o minte), credinţele şi raţionalitatea (sau lipsa acesteia) credinţelor. O voi folosi în esenţă în acest mod, deşi adesea voi include cunoaşterea şi cunoaşterea tuturor speciilor sub o umbrelă filosofică rezervată în mod normal fiinţelor umane. Pe scurt, fiţi pregătiţi pentru afirmaţii despre „natura epistemică a minţii unei bacterii”. Existenţial se referă fie la existenţă, fie la şcoala de gândire cunoscută sub numele de existenţialism. Îl voi folosi doar în primul sens şi numai pentru a mă referi la entităţi (reale sau teoretice) despre care putem concluziona cu uşurinţă că există, sunt reale şi au statut ontologic. Doresc să îmi prezint argumentele într-un mod echilibrat, să fiu clar, dar nu obtuz, să fiu precis, fără a călca pe urme lexicografice. Atunci când voi aborda alte aspecte filosofice (pe care le vom aborda, ele sunt inevitabile), cum ar fi dualismul, materialismul, funcţionalismul, realismul, misterianismul, panpsihismul şi reducţionismul, voi face tot posibilul să fiu clar şi să prezint diferitele moduri în care sunt utilizate.

Există o metodă în nebunia mea. Impunerea deliberată a alunecării denotative nu se datorează faptului că sunt leneş (deşi sunt). Este menită să netezească calea pentru dezvoltarea modelului **CBC** - şi anume, că toate aceste stări ontologice, care apar de la un capăt la celălalt al continuumului filogenetic, indiferent despre ce specii vorbim sau de modul în care acestea diferă unele de altele în ceea ce priveşte forma fizică sau sofisticarea cognitivă, sunt subiective, experimentate intern şi au o bază comună, cauzală, biologică. Comuniunea şi continuitatea sunt temele de aici.

O paranteză privind natura argumentului filozofic

Deşi nu am o pregătire formală în filosofie, am ajuns să ştiu câte ceva. Am urmat un singur curs în domeniu; se intitula „Filosofia istoriei’’, şi m-am înscris la el în primul rând pentru că nu-mi

puteam imagina cum ar putea fi așa. Nu se întâmpla istoria pur și simplu? Cum ar putea exista o filosofie a ei? Ei bine, a existat/există, dar cursul m-a făcut să mă interesez de istorie, iar istoria psihologiei a devenit una dintre subspecialitățile mele.

Două decenii mai târziu, în 1981, am avut norocul de a fi unul dintre puținii psihologi cognitivi (dacă memoria nu mă înșeală, eram patru) invitați să participe la un institut de vară de șase săptămâni privind psihologia și filosofia minții, sponsorizat de *National Endowment for the Humanities*. Ni s-a alăturat un număr mult mai mare de filosofi (îmi amintesc că au fost cam 40) în campusul Universității Washington din Seattle.

Acea vară a stârnit în mine un interes profund pentru filozofie și, din fericire, acesta a apărut fără capcanele obișnuite ale formării formale. Spun „din fericire" pentru că, pe parcursul acelor săptămâni de interacțiune intensă cu persoane - dintre care multe erau sau au devenit cercetători importanți[8] și cu mulți dintre ei m-am împrietenit și pe care încă îi consider prieteni - am fost capabil să absorb idei din mers, să las procesele osmotice intelectuale să mă umple de înțelegere.[9] Pentru că nu aveam o pregătire formală, am descoperit că puteam să citesc literatura de specialitate, să dialoghez cu cei care erau pătrunși de tradițiile domeniului, dar să fac acest lucru într-un mod relativ liber. De multe ori m-am trezit oferind perspective noi sau critici care nu fuseseră luate în considerare înainte. A existat și un dezavantaj. Recunosc că am făcut multe greșeli atunci când, de exemplu, m-am împiedicat de o problemă care fusese deja analizată în detaliu - deși nu cunoșteam soluția - și adesea am folosit greșit un termen sau am înțeles greșit o teorie. Dar, pe termen lung, autonomia ignoranței a fost un bonus. Sunt destul de sigur că nu aș fi putut scrie niciodată această carte dacă nu aș fi avut acele șase săptămâni în Seattle, sponsorizate de NEH. Sunt la fel de sigur că nu aș fi putut scrie niciodată această carte dacă aș fi avut o pregătire formală în filosofia minții.

La sfârșitul institutului, am plecat și cu o noțiune conspirativă, vicleană. Am ajuns să bănuiesc că atunci când un filozof își ia doctoratul, semnează un jurământ secret care sună cam așa: „*Promit să nu fiu niciodată de acord cu niciunul dintre colegii mei. Le voi critica teoriile și le voi ataca interpretările. Îmi voi apăra poziția cu toate argumentele nuanțate posibile. Dacă mă răzgândesc, o voi face într-un mod obscur, astfel încât să nu fie evident pentru nimeni. În cazul în care toate acestea eșuează, voi inventa o nouă versiune a oricărui principiu luat în considerare și îi voi da un nume improbabil.*"

Astfel, practic fiecare principiu sau perspectivă care merită atenția filosofilor are mai multe variații nuanțate și, deși mă distrez puțin pe seama colegilor mei, în cea mai mare parte, aceștia sunt conștiincioși în ceea ce privește distincțiile pe care le fac și le descoperă pe cele care merită luate în considerare. După cum am sugerat mai devreme în paragrafele referitoare la terminologie, multe dintre problemele pe care le vom analiza în acest mic volum sunt cele pe care filosofii le-au mestecat, unii dintre ei timp de secole și, în acest fel, au crescut mai multe ramuri conceptuale. Poziții precum *dualismul* au început prin a propune existența a două domenii ontologice, unul fundamentat pe materiale fizice, fapte și comportamente observabile și altul lipsit de substanță, dar care se dezvăluie în alte moduri. Acest lucru părea simplu și, într-adevăr, primele versiuni au apărut din considerente teologice și împărțeau existența umană într-un corp material și un suflet eteric care transcendea forma fizică. Din păcate, de atunci, o întreagă familie de dualisme a apărut în domeniu, inclusiv dualismul proprietății, dualismul predicatului și dualismul substanței. Există variații pe fiecare dintre aceste teme, modalități multiple de a aborda fiecare dintre aceste subcategorii ale dualismului. Se poate intra în luptă dintr-un unghi interacționist, epifenomenalist sau paralelist.

---

8 Au participat în calitate de invitați, conducători de discuții sau invitați John Searle, Jerry Fodor, Ned Block, Owen Flanagan, Bill Bechtel, Carol Cleland, Fred Dretske, Barry Loewer, Zenon Pylyshyn, Jonathan Adler, Paul Thagard, Lance Rips, Bill Brewer, Stu Silvers, Bob McCauley, Dan Dennett, Danny Kahneman, Georges Rey, Barbara van Eckardt, Stephen Stitch și John Heil - un grup cel puțin remarcabil, iar această listă îi include doar pe cei pe care mi i-am putut aminti sau urmări.

9 În mod interesant, sociologul Harry Collins de la Universitatea Cardiff a explorat modul în care cineva poate ajunge să dobândească ceea ce el numește „expertiză interacțională" într-un domeniu nou prin simpla imersiune în lumea socială a celor din domeniul respectiv. A se vedea Collins & Evans (2015) pentru detalii. Cred că am experimentat ceva de genul acesta în acea vară.

Funcționalismul a început ca o poziție care sublinia importanța fundamentării analizei pe funcțiile pe care le îndeplinește un sistem sau o entitate, mai degrabă decât să fie distrasă de o analiză excesivă a formei sau structurii sale. Pornind de la această bază simplă, au apărut o jumătate de duzină sau mai multe variante, inclusiv funcționalismul computațional, funcționalismul stării mașinii, funcționalismul independent de hardware, funcționalismul analitic, funcționalismul rolului și al realizatorului și o duzină de altele.

Emergentismul a fost introdus pentru prima dată într-un mod destul de precis pentru a identifica o enigmă: cum pot anumite proprietăți sau funcții să apară sau să se nască din relațiile structurale sau interacțiunile dintre formele materiale? În literatura de specialitate există în prezent atât de multe forme distincte încât, în cartea sa din 2007, *Mind in Life*, biofilosoful Evan Thompson s-a simțit obligat să adauge un apendice de 15 pagini în care să analizeze varietatea de poziții emergentiste susținute în prezent.

Acest model de proliferare a pozițiilor poate fi găsit în aproape fiecare domeniu al filosofiei - și acest lucru este în regulă, este OK. Este treaba filosofilor să ne mențină cinstiți pe noi ceilalți, să supună teoriile, paradigmele și interpretările la teste dure, să ascută argumentele și afirmațiile pe o roată aspră. Și, pentru că acest volum va păși pe terenuri filosofice, va trebui să mă confrunt cu unele dintre acestea. Când o voi face, voi folosi aceeași abordare dură ca și în cazul terminologiei - în cea mai mare parte. *Dualismul* va fi abordat ca o poziție singulară care susține că există două tipuri de „lucruri" acolo, lucruri fizice și lucruri nonfizice (deși unii filosofi s-ar putea să se strâmbe la folosirea cuvântului „lucruri"). *Emergentismul* se va referi la o problemă întâlnită în mod regulat de oamenii de știință: Cum apare sau se exprimă o funcție sau o proprietate a unui sistem? De ce apa este umedă atunci când nu există nici oxigen, nici hidrogen? Cum creează sistemul vizual senzația de *roșu* atunci când nu există „roșu", ci doar suprafețe care fie absorb lumina de lungime de undă scurtă și medie, fie emit direct radiații electromagnetice în apropiere de 700 nm?

Atunci când este necesar, vor fi introduse mai multe detalii - uneori în text, alteori într-o notă de subsol. Ideea este de a încerca să rămânem în limitele bunului simț, fără a le face rău specialiștilor. O mare parte a dezbaterilor asupra acestor și altor probleme din filosofie forțează aceste limite. Exercițiile pot fi interesante și, ocazional, iluminatoare din punct de vedere filosofic, dar sunt adesea distrageri ale atenției. Eu nu vreau să fiu distras.

Distincția între biologia evoluționistă și psihologia evoluționistă

Teza **CBC** se bazează pe principiile de bază ale biologiei evolutive și diferă, în abordarea sa metodologică generală, de cea a subspecialității din cadrul științelor sociale cunoscută sub numele de *psihologie evolutivă*. Doresc să fac această distincție de la bun început. Am constatat că mulți nu înțeleg că acestea sunt două subdomenii diferite de cercetare în cadrul domeniului mai larg al teoriei evoluționiste. De asemenea, am avut mai multe ocazii în care cititorii au confundat abordarea mea ca fiind una care ține de psihologia evoluționistă. Nu este așa.

Deși ambele discipline utilizează principiile fundamentale ale evoluției ca parte a suportului lor teoretic și conceptual, ele folosesc metodologii diferite și examinează probleme diferite. *Psihologia evoluționistă* analizează rolul adaptiv pe care anumite mecanisme psihologice și comportamentale ipotezate l-ar fi putut avea în evoluție și le utilizează ca puncte de plecare pentru a explora comportamentul uman (și, ocazional, al altor specii). Studiile clasice sunt cele care au explorat diferențele de sex folosind teoria investiției sexuale a lui Robert Trivers, argumentând că femelele, deoarece investesc mai mult timp și efort în reproducere (gestație, naștere, alăptare), vor fi mai discriminatorii în selecția partenerului decât masculii - o constatare larg susținută în aproape toate culturile. Alte studii stabilesc o legătură între factori precum apropierea de rude sau gene comune între indivizi și manifestările de altruism și comportamentul de ajutorare. Domeniul are o

arie de cuprindere extinsă într-o serie de comportamente - prezentarea generală a lui David Buss din 2011 este un loc bun de vizitat pentru oricine este interesat.

Abordarea pe care o adopt aici (şi pe care am adoptat-o în trecut în analiza mea a *inconştientului cognitiv* în Reber, 1989 şi 1993) diferă atât în ceea ce priveşte metoda, cât şi modul de teoretizare. Poziţia **CBC** utilizează principiile fundamentale ale biologiei evolutive pentru a face predicţii cu privire la tipurile de forme, funcţii şi comportamente la care ar trebui să ne aşteptăm în anumite specii. Să luăm, de exemplu, principiul fundamental al biologiei evolutive conform căruia, odată ce o trăsătură sau o funcţie capătă un avantaj adaptiv, aceasta tinde să se menţină în cadrul speciei şi, prin extensie, în cadrul tuturor speciilor care au evoluat din aceasta. Pe scurt, acest principiu afirmă că evoluţia este conservatoare, iar pierderea unei caracteristici adaptative la o specie care a evoluat de la una care o prezenta ar trebui să fie un eveniment extrem de rar. Un alt principiu este cel al stabilităţii, care afirmă că, odată ce o trăsătură sau o structură s-a stabilit într-o specie şi alte trăsături au evoluat pe baza ei, aceasta devine stabilă, mai puţin susceptibilă de a suferi mutaţii şi va prezenta mai puţine diferenţe individuale. Un bun exemplu care reflectă ambele principii este vertebra. Coloana vertebrală are o valoare adaptivă semnificativă şi, prin urmare, odată stabilită, a devenit baza pentru toate speciile de vertebrate. Aceasta prezintă relativ puţine variaţii, păstrându-şi forma fundamentală într-o gamă largă de specii.

Un mod de a vedea diferenţa dintre abordarea pe care o adopt aici şi *psihologia evoluţionistă* este să observăm că aceasta din urmă priveşte „înapoi”, examinând posibilele rădăcini genetice ale comportamentelor observate în prezent, în timp ce biologia evoluţionistă priveşte „înainte”, explorând modul în care trăsăturile au evoluat din forme anterioare pe baza unor principii euristice fundamentale. Mai târziu, în capitolul 5, vom analiza mai atent modul în care teoria **CBC** se încadrează în acest tip de cadru evolutiv. Precizez această distincţie aici doar pentru a preveni orice confuzie. S-a mai întâmplat.

Biologia evoluţionistă a fost întotdeauna în centrul conceptual al cercetării mele. Este singurul cadru teoretic mai larg în care m-am simţit confortabil să lucrez. Una dintre maximele directoare este că orice mecanism sau trăsătură luată în considerare trebuie să fie una care are sens în cadrul biologiei evolutive. Această doctrină este, de exemplu, unul dintre motivele pentru care nu am fost niciodată mulţumit de cadrul lingvistic prezentat de Noam Chomsky (a se vedea Reber, 2011). Ipoteza unui „organ” lingvistic complet echipat cu o formă de codificare genetică a unei gramatici universale pur şi simplu nu avea sens din perspectiva darwinistă. În timp, susţinătorii poziţiei de tip Chomsky au ajuns să recunoască acest lucru - inclusiv Chomsky - şi au fost făcute o serie de ajustări ale teoriei (a se vedea Fitch. Hauser, & Chomsky, 2005).

O notă privind citatele

Majoritatea materialelor vor fi prezentate fără citare, cu excepţia menţionării numelui omului de ştiinţă sau filosofului care a fost/este principalul responsabil pentru problema în discuţie. Lucrarea autorului despre care se discută poate fi găsită în lista de referinţe sub numele persoanei respective, cu o singură excepţie. Există o diferenţă procedurală curioasă în ceea ce priveşte modul în care psihologii şi filosofii enumeră autorii unei lucrări în colaborare, comparativ cu cercetătorii din domeniul ştiinţelor biomedicale. În psihologie şi filosofie, procedura standard privind ordinea autorilor într-o lucrare este ca numele autorului principal să fie primul. Prin urmare, atunci când mă refer la un articol, de exemplu, la „Jones” şi ea/el a fost, de fapt, autorul principal, nu mă voi obosi să dau nicio informaţie suplimentară. Lucrarea poate fi găsită în lista de referinţe sub „Jones” pentru oricine este interesat să urmărească materialele originale.

Ştiinţele biomedicale utilizează o procedură diferită. Acolo, protocolul standard este ca şeful laboratorului în care a fost efectuată cercetarea să îşi treacă numele la sfârşit, primul autor fiind omul de ştiinţă care a avut un rol esenţial în proiectul respectiv. Prin urmare, atunci când discut despre munca cuiva, să spunem profesorul „Smith”, care conduce un astfel de laborator şi

supraveghează toate eforturile acestuia, mă voi referi la descoperirile lui „Smith”, dar, desigur, lucrarea relevantă nu va fi găsită sub numele său. În aceste cazuri, voi cita lucrarea după numele primului autor, astfel încât textul, de exemplu, va suna astfel: „Laboratorul lui Smith a descoperit blah blah blah. Consultați Anderson et al. (2017) pentru detalii”.

Rezumat

Scopul meu în acest mic volum va fi să explic modelul **CBC**, să explic de ce are sens, de ce alte abordări ale originilor conștiinței și minții sunt fie fatalmente greșite, fie întâmpină dificultăți serioase și poate insurmontabile. Ne vom uita la datele care există pentru a susține ipoteza de bază și, ceea ce este important, care sunt implicațiile modelului. Abordarea **CBC**, susțin eu, oferă o strategie optimă de cercetare tocmai pentru că invită la reconfigurarea multor aspecte din biologia evolutivă, microbiologie, științele cognitive și, da, din filosofia minții.

Mulțumiri

Mulțumesc colegilor și prietenilor care au încercat să mă salveze de mine însumi, în special lui Janet Atkinson-Grosjean, Frantisek Baluska, Dennis Bray, Harry Collins, Antonio Damasio, Daniel Dennett, Stevan Harnad, Preston Oade, Carolyn Ristau și grupului obișnuit de recenzenți anonimi. O mulțumire specială lui Dennis Bray, în primul rând pentru cartea sa *Wetware* despre minunata lume a ființelor vii foarte mici și, în al doilea rând, pentru feedback-ul său blând și corectiv asupra materialului din capitolul 4.1 m-am aventurat în domeniul microbiologiei ca un adevărat naiv. Am găsit cartea lui Bray din 2009 și a devenit ghidul meu prin lumea organismelor unicelulare. Îi mulțumesc pentru că a scris atât de bine și de convingător despre lumea uimitoare a acestor forme de viață foarte mici și uimitoare.

Mulțumesc, de asemenea, editorilor mei de la *Oxford University Press*, Phil Velinov și Joan Bossert. O cunosc pe Joan de mai mult timp decât ar vrea oricare dintre noi să recunoască. Ea a editat prima mea carte *OUP, Implicit Learning and Tacit Knowledge*, publicată în 1993 (și, în mod uimitor, încă tipărită un sfert de secol mai târziu). Joan, pe lângă faptul că este unul dintre cei mai eficienți, sensibili și perspicace editori din lumea publicațiilor academice, este o mare gurmandă; unul dintre dezavantajele mutării mele în Pacific Northwest în 2005 a fost acela de a nu putea merge cu ea la vreun restaurant minunat din New York în timp ce acest proiect continua.

De asemenea, îi mulțumesc în mod special lui Rhiannon Allen pentru sfaturile, feedback-ul, editarea și pentru că a suportat toate hărmălaia, discuțiile, întrebările și insistențele cu privire la probleme, preocupări, nedumeriri și spasme de frustrare atunci când pur și simplu nu știam despre ce este sau despre ce ar putea fi vorba. Până acum, ea s-a obișnuit cu acest lucru, versiunea academică a abuzului conjugal, și reușește să ofere o contribuție pătrunzătoare și adesea convingătoare. Va nega că ar fi avut vreun rol aici, dar asta poate pentru că nu înțelege cât de convingător este să expui o idee și să urmărești și să asculți cum reacționează. O sprânceană ridicată sau un cap înclinat sunt suficiente pentru a-mi spune că trebuie să mă întorc și să mă gândesc mai profund.

Bacteria captivantă care funcționează ca un marcator de paragraf provine din lucrarea lui Max Verworn și Frederic Lee, *General Physiology: An Outline of the Science of Life* (1899); Londra, New York: Macmillan and Co. Rhiannon l-a găsit și l-a foto-shopat într-o formă utilizabilă.

Point Roberts, Washington
și
Departamentul de Psihologie Universitatea din British Columbia
Vancouver, BC

1. Pot roboții să fie conștienți?

Înainte de a dezvolta mai detaliat teoria *bazei celulare a conştiinţei* (**CBC**), trebuie să fac o trecere la o problemă care, de foarte mult timp, a fost atât seducătoare, cât şi controversată: trebuie ca acest lucru al conştiinţei, acest lucru al minţii să fie o proprietate a unei fiinţe umane? Sau, dacă nu este un om, poate un alt animal destul de complex şi foarte funcţional, cum ar fi un *bonobo* sau o cioară? Poate un delfin sau o caracatiţă? Este adevărat că minţile pot locui doar în organisme organice, bazate pe carbon? Sau, concentrându-ne asupra întrebării care se pune cu adevărat, ar putea o entitate artificială care să aibă aceleaşi comportamente observabile, să se angajeze în aceleaşi operaţiuni de calcul ca oamenii (sau o altă specie) să aibă, de asemenea, stări mentale, minți, conştiinţă? Ar putea o inteligență artificială (AI) să devină conştientă, cu adevărat conştientă de sine, de ceea ce o înconjoară, de ceilalți? Și cum am putea ști?

Funcționalismul sau posibilitatea conștiinței „independente de hardware

Această problemă a fost, până de curând, unul dintre subiectele fierbinți în rândul filosofilor minții, în special al celor interesați de cercetările efectuate în domeniul roboticii şi al neuroştiinţelor cognitive. Deşi în secolul al XVII-lea Thomas Hobbes a cochetat cu o noțiune mecanicistă similară (conform căreia esenţa activităţii minţii umane este măsurarea şi calcularea, funcţii care, în principiu, pot fi îndeplinite de alte entități sau sisteme), primele licăriri coerente a ceea ce a ajuns să fie cunoscut sub numele de teza **IA puternică** au apărut în anul 1961 într-o povestire scurtă, „*Dial F for Frankenstein*”, scrisă de marele scriitor de *science fiction* Arthur C. Clarke. În lumea fantastică pre-Internet creată de el, sistemul telefonic a devenit din ce în ce mai complex și mai interconectat, pe măsură ce tot mai mulți oameni, corporații și organizații au fost conectați la rețea. În cele din urmă, rețeaua telefonică mondială a atins masa critică și a trecut un prag metafizic pentru a deveni conştientă. Această minte nouă, foarte artificială, și-a anunțat sosirea cu un strigăt de naștere - toate telefoanele din lume au sunat simultan și haosul s-a instalat rapid, pe măsură ce acest nou monstru „Frankenstein” își exercita puterea. De altfel, Sir Timothy Berners-Lee, unul dintre informaticienii care au contribuit la dezvoltarea *World Wide Web*, l-a creditat pe Clarke pentru inspirație, menționând că, citind această poveste, a realizat pentru prima dată că ceva asemănător internetului modern era posibil.

Tot în 1961, renumitul filosof Hilary Putnam a prezentat primele argumente logice în favoarea propunerii conform căreia o entitate artificială, independentă de caracteristicile sale fizice, ar putea fi echivalentă din punct de vedere cauzal cu un creier uman. Putnam a argumentat că, dacă o inteligență artificială instanțiată pe un computer digital ar efectua operațiuni identice cu cele ale unui creier uman, proprietățile sale cauzale și, prin urmare, stările sale subiective interne ar fi echivalente din punct de vedere funcțional cu creierul respectiv. Putnam a denumit această poziție „funcţionalismul stării maşinii”. Propunerea a fost preluată de o serie de filosofi și de oameni de știință cognitivi și informaticieni, printre care Douglas Hofstadter, Daniel Dennett, David Chalmers și Jerry Fodor, student al lui Putnam. După cum vom vedea, această abordare și-a pierdut mult din strălucire din cauza unor critici viguroase şi convingătoare din diverse surse, cum ar fi lucrările importante ale filosofului John Searle. Dar ea încă răsună puternic și seducător în lumea futuriștilor, în *science fiction* și în rândul publicului larg. Ideea că mintea sau conştiinţa pot exista în operaţii şi calcule, indiferent de tipurile de entități care le efectuează, are ceva atrăgător. Iar întrebările-cheie care planează asupra tuturor acestor aspecte sunt: „Cum am putea ști?” „*Cum putem ști dacă ceva sau cineva este conştient, are o minte, este sensibil?*”

Există motive întemeiate pentru care aceste întrebări continuă să apară. În primul rând, trebuie să recunoaștem că suntem blocați în perspectiva noastră la persoana întâi, în faptul că noi,

tu, eu, noi toți, ne cunoaștem cu adevărat doar propriile minți. Cum știm, de exemplu, dacă un cimpanzeu are o formă de conștiință? Dar o sepie? Are gânduri? Sau pisica mea. Chiar înțelege ceva? Îmi place să răspund afirmativ la fiecare dintre aceste întrebări și, de fapt, sunt destul de încrezător că toate au minte. Întreaga schelă pentru modelul meu **CBC** este construită pe acest principiu fundamental. Dar haideți să împingem puțin limitele, ce ziceți de un robot (sau *hot*, în jargonul curent)? Are sau ar putea ajunge să aibă o formă de conștiință, așa cum a sugerat Putnam? Putem spune că Watson, supercomputerul care, în 2011, i-a învins fără probleme pe Ken Jennings și Brad Rutter, cei doi campioni umani la Jeopardy, într-un meci de poveste, are gânduri autentice? La urma urmei, Watson părea să știe aproape totul și a răspuns la întrebări[10] cu o viteză uimitoare, pe care nici Jennings, nici Rutter nu au putut-o egala. Și, desigur, ce se va întâmpla cu o viitoare entitate artificială care chiar va trece testul Turing[11] sau care va vorbi gramatical engleza sau chineza?

Punctul meu de vedere este de cealaltă parte a acestei diviziuni ontologice. Nu, nu cred că un robot are conștiință sau ar putea fi conștient. Nu, nu cred că Watson știe ceva. Și nu, nici nu aș ajunge la concluzia că un computer care ar trece testul Turing ar ști ceva - și niciunul nu s-a apropiat nici măcar atât timp cât gama de categorii permise a fi introduse nu este restricționată într-un fel[12]. Dar de ce? Simpla exprimare a convingerilor mele, chiar și atunci când acestea sunt larg împărtășite, nu va face fericit niciun filosof care să merite. Avem nevoie de argumente. Să începem cu ceea ce pare a fi o întrebare destul de simplă - poate părea o plimbare pe o cale fără subiect, dar există un motiv pentru aceasta.

Ce știu eu/voi despre minți - a ta și/sau a mea?

Ei bine, în primul rând, știu că am o minte și bănuiesc că și tu simți cam același lucru despre tine. De la Descartes încoace, această propunere foarte personală este cea mai frecvent citată drept trăsătură definitorie a ceea ce înseamnă să fii om. Dar, așa cum am menționat mai sus, de unde știu eu că tu ai o minte? O manevră filosofică populară este experimentul de gândire cunoscut sub numele de argumentul *zombie* - uneori numit argumentul filosofic *zombie* sau argumentul *p-zombie*. Cuvântul „filosofic” a fost introdus pentru a distinge acești *zombi* de strigoii din folclorul haitian. Un *zombi* arată exact ca o ființă umană, se comportă ca o ființă umană și răspunde la întrebări ca o ființă umană, dar îi lipsesc stările subiective, *qualia*[13], conștiința. Toate acțiunile și declarațiile sale sunt simulări ale comportamentelor și vorbelor umane. Dacă lovești cu pumnul un *zombi*, acesta se va retrage de la lovitură spunând ceva de genul „Hei, asta a durut”, deși nu simte nicio durere reală.

10 Pentru cei care nu sunt familiarizați cu popularul joc de televiziune Jeopardy, răspunsurile sunt furnizate (de exemplu, „Capitala Franței”), iar concurentul trebuie să răspundă cu întrebarea („Ce este Parisul?”). În cazul în care răspunsul nu este formulat în acest mod, concurentul pierde premiul pentru itemul respectiv, chiar dacă acesta este corect din punct de vedere faptic.

11 În acest test, făcut celebru de matematicianul Alan Turing în 1950, un computer se află într-o cameră, iar o persoană în cealaltă. Un evaluator pune întrebări ambilor, folosind limbajul natural, pe orice subiect dorește. Dacă evaluatorul nu poate spune cine (sau ce) se află în fiecare cameră, se poate spune că entitatea artificială a dat dovadă de inteligență umană reală și a „trecut testul Turing”. Rețineți că Turing nu s-a aventurat să se întrebe dacă un astfel de computer de succes ar avea „conștiință”, ci doar dacă a demonstrat „inteligență”.

12 În fiecare an se organizează așa-numitul concurs Loebner Prize pentru IA care obține cele mai bune rezultate la un test Turing controlat. Concursul a fost inițiat în 1990 de controversatul inventator american Hugh Loebner împreună cu Centrul Cambridge pentru Studii Comportamentale. Acesta a fost lăudat de mulți ca încurajând cercetarea în domeniul inteligenței artificiale și criticat vehement, în special de către cercetătorul Marvin Minsky, ca fiind o cascadorie publicitară care nu face nimic pentru a promova progresul real în domeniul inteligenței artificiale.

13 Qualia (sing, quale) sunt experiențe subiective, senzoriale. Qualia au un element „cum este”, cum ar fi cum este să mușți dintr-un ardei iute thailandez sau să auzi un tenor atingând un Do înalt sau să simți durerea unei tăieturi de cuțit. Există mai mult decât câteva dezbateri ezoterice în filosofie cu privire la natura, rolurile cauzale și chiar existența qualia, dar pentru scopurile noastre gândiți-vă doar la ele ca la experiențe senzoriale și totul va fi bine.

Dacă îl întrebi dacă este conştient sau are o minte, va minţi printre dinţii lui de *zombi* şi va răspunde: „Da, desigur, la fel ca tine”.

Zombi-ul a fost introdus de filosofi precum David Chalmers[14] într-un efort de a critica argumentele neo-comportamentaliste şi materialiste prezentate de filosofi precum Daniel Dennett, care, în cartea sa din 1991, *Consciousness Explained* (Conştiinţa explicată), a afirmat că factorii cauzali din minţile noastre ar putea fi descrişi pe deplin prin acţiunile, declaraţiile şi experienţele noastre, fără a fi nevoie să se formuleze ipoteza unei entităţi ontologice separate, precum conştiinţa.[15] Dennett nu susţine că nu există ceva precum experienţa subiectivă care să se potrivească sensului psihologiei populare al conştiinţei. El susţine că aceasta nu are niciun rol cauzal independent de jucat. Este, în termenii săi, o „iluzie a utilizatorului” care rezultă din faptul că creierul gândeşte, decide, devine conştient de mişcare, îşi aminteşte lucruri, toate acestea având loc în mod întâmplător şi paralel.

Dennett susţine că, dacă ajungem să înţelegem care sunt aceste procese bio-neurale şi modul în care fiecare este legat cauzal de o anumită emoţie, acţiune, imagine, intenţie, gând sau orice altceva care se întâmplă în capul nostru, vom înţelege mintea - şi vom face acest lucru fără a invoca conceptul unui agent ontologic independent pe care alţii îl numesc „conştiinţă”.

Din punctul de vedere al lui Chalmers, linia de raţionament a lui Dennett are o problemă deoarece, după cum susţine acesta, având în vedere că *zombi*-ul nostru ipotetic este echivalent fiziologic cu un om, dar nu are stări mentale, nu are conştiinţă, conştiinţa noastră trebuie să derive din factori nemateriali. Faptul că putem conceptualiza posibilitatea unui *p-zombie*, a concluzionat Chalmers, înseamnă că ar putea exista şi că, prin urmare, conştiinţa umană poate fi înţeleasă pe deplin doar într-un cadru dualist în care minţile nu sunt echivalente din punct de vedere cauzal cu funcţiile neuronale ale creierului - este nevoie de altceva.

Refutarea standard a materialistului[16] la scenariul *zombi* este de a sublinia că, deşi un *zombi* ar putea fi o posibilitate conceptuală, este aproape sigur o imposibilitate metafizică. Având în vedere că poziţia materialistului se bazează pe presupunerea că doar structurile şi funcţiile fiziologice şi neurologice sunt cele care produc conştiinţa, orice efort de a crea un adevărat *zombi* ar eşua pur şi simplu pentru că acesta ar trebui să aibă stări mentale, *qualia* şi conştiinţă şi, prin urmare, să nu fie deloc un *zombi*. Există şi o replică la această afirmaţie, dar nu vă voi plictisi cu ea. Pentru cei interesaţi, Robert Kirk are o prezentare foarte completă a discuţiilor despre *zombi* de toate tipurile (a se vedea Kirk, 2003, revizuit în 2015).

Dar, indiferent de partea în care vă situaţi în această dezbatere, ideea rămâne valabilă: Cum pot să ştiu, să ştiu cu adevărat, că aveţi o minte precum cea pe care ştiu eu că o am? Problema, spune argumentul standard al „altor minţi”, este că avem doar o perspectivă la persoana întâi. Nu cunoaştem alte minţi în afară de a noastră şi, prin urmare, toate datele despre alte minţi provin în mod necesar dintr-o perspectivă de a treia persoană. Noi doar ne uităm la celălalt şi facem presupuneri.

Cel mai simplu răspuns, cel din psihologia populară, care ocoleşte *zombii* şi alte arcanuri filosofice, este că, deoarece ştiu că am o minte, este nevoie doar de o măsură modestă de generozitate din partea mea pentru a vă acorda aceleaşi tipuri de experienţe fenomenale pe care le am eu. În lucrarea lor fundamentală din 1978, David Premack şi Guy Woodruff au formulat această problemă succint. Fiecare dintre noi are o teorie a minţii, pe care am dezvoltat-o în copilărie din interacţiunile sociale cu ceilalţi şi, odată ce am dobândit această înţelegere, putem aprecia că fiecare

14 Noţiunea de automat viu care se comportă, dar nu simte datează încă de la Descartes, care privea toate formele de viaţă non-umane exact în acest fel. Noţiunea de zombie ipotetic a fost abordată şi de filosofi precum Saul Kripke şi Ned Block, dar Chalmers a fost cel care a dus lupta direct la materialişti.

15 Criticii tezei lui Dennett s-au distrat copios redenumindu-i cartea. David Barash a sugerat *Consciousness Explained Away*, iar Peter Carruthers a propus *Consciousness Ignored*.

16 În conformitate cu lexicografia mea lejeră, voi folosi materialismul, fizicalismul şi termenii derivaţi ca sinonime aproximative care identifică o poziţie care susţine că toate fenomenele cunoscute, atât fizice, cât şi mentale, pot fi pe deplin surprinse de principii care implică materia biofizică şi argumentul mecanicist. Reţineţi, de asemenea, că fizicalismul este strâns legat de reducţionism, principiul conform căruia principiile explicative optime trebuie găsite prin explorarea elementelor cauzale subiacente mai profunde din natură. Voi avea mai multe de spus despre reducţionism mai târziu, deoarece **CBC** este o teorie extrem de reducţionistă.

are nevoi și dorințe, dureri și temeri, intenții și gânduri la fel ca ale noastre și, ceea ce este important, că putem folosi aceste informații pentru a evalua, interpreta și face predicții cu privire la ceea ce vor face. Vă recunosc ca fiind un conspecific, un membru al speciei mele *Homo sapiens*, și mi se pare firesc și corect să vă înzestrez cu ceea ce am eu. Și, din considerente strict pragmatice, va fi mult mai ușor să facem știință a conștiinței dacă pornim de la premisa că împărtășim un set comun de stări interne și experiențe fenomenale - chiar dacă suntem încătușați de perspectiva noastră la persoana întâi și nu putem ști cu siguranță ce experimentează alții.

Când vine vorba de alte specii, cimpanzei, pisici, caracatițe, lucrurile pot deveni puțin mai dificile și sentimentul de înțelegere începe să se destrame. În primul rând, începem să pierdem euristica conspecifică. Cimpanzeii sunt, ca și noi, primate. Deci, bine, este suficient de aproape pentru mulți. Pisicile cu siguranță nu sunt primate, dar sunt mamifere și uneori se comportă ca și cum ar avea stări mentale, în special atunci când fac genul de lucruri diabolice pentru care sunt renumite pe bună dreptate. Deci, da, suficient de bine. Cum rămâne cu caracatițele? Se știe că caracatițele ținute în captivitate fac lucruri precum spargerea becurilor enervant de luminoase prin stropirea lor cu apă, ceea ce a impresionat nu puțini dintre cei care studiază cefalopodele. În cartea sa *Other Minds*, biofilosoful Peter Godfrey-Smith concluzionează că primele licăriri evolutive ale unei conștiințe autentice au apărut probabil la caracatițe și la sepie.

Dar observați ce se întâmplă aici. Cu cât ne îndepărtăm mai mult de noi, cu atât mai fragil devine argumentul din perspectiva terților, cu atât mai mult ne confruntăm cu o altă problemă. Chiar dacă ne extindem simțul generozității fenomenale, subiective la aceste alte specii, tot nu știm dacă vreuna dintre aceste creaturi are o viață mentală care se află pe un continuum cu a noastră. Dacă suntem presați de acest lucru, ne putem baza întotdeauna pe lucruri precum familiaritatea antropomorfă. Cimpanzeii seamănă destul de mult cu noi, sunt cunoscuți ca fiind creatori de unelte, au mecanisme bogate de învățare socială și manifestă înclinații culturale clare cu caracteristici comune cu copiii umani (a se vedea Whiten. 2017 pentru o analiză bună).

Astfel, noțiunea de minte a cimpanzeilor este destul de ușor de recunoscut pentru majoritatea oamenilor. De fapt, în lucrarea lui Premack și Woodruff care a introdus noțiunea de teorie a minții, întrebarea se referea în special la cimpanzei.

Atunci când Obadiah, una dintre pisicile noastre, se ghemuiește în spatele bibliotecii și sare la mine când vin după colț (și o face) sau când se apropie de unul dintre noi, miaună și apoi pleacă de parcă ar cere o partidă bună de „urmărește o pisică” (și o face), mă simt confortabil să îi atribui stări mentale, speranțe, dorințe și o anumită măsură de capacitate de rezolvare a problemelor. Pentru că nu pot folosi stratagema aproape specifică aici, trebuie să folosesc un alt tip de justificare internă pentru această mișcare. Ceea ce fac este să împing un pic unghiul antropomorfismului. Presupun un punct comun evolutiv cu Obi, deoarece comportamentele sale îmi amintesc de lucrurile pe care le-am făcut, mai ales când eram copil, pentru a-i determina pe alții să joace un joc. Desigur, nu toată lumea se simte confortabil cu această mișcare și nu toți cei care au abordat acest subiect și altele conexe i-ar conferi lui Obi ceva apropiat de tipul de viață internă și mentală pe care tocmai am făcut-o eu. Cu cefalopodele, manevra începe să pară tensionată. Da, caracatițele pot fi, așa cum am menționat, soluționatoare inteligente de probleme. Ele au, de asemenea, sisteme nervoase bogate și complexe, construite pe o schelă complet diferită de cea a vertebratelor.[17] Dar bănuiesc că persoana obișnuită de pe stradă se va opune acordării unei minți unui cefalopod.

În aceste cazuri și în multe altele, în care ne întrebăm dacă sau cum putem să interpunem noțiunile noastre foarte umane de minte în alte creaturi și dacă sau cum putem fi vreodată încrezători în concluziile noastre, ne îmbrățișăm „Nagel-ul interior”. În lucrarea din 1974 menționată mai devreme, Thomas Nagel s-a confruntat direct cu această problemă a altor minți într-una dintre cele mai celebre lucrări de filosofie intitulată „*Cum este să fii liliac?*” Răspunsul lui Nagel a fost pur și simplu că nu știm. De fapt, nu putem ști pentru că suntem legați de perspectiva noastră la persoana întâi. Putem să presupunem cât vrem, dar singura creatură care știe cum este să

17 Căile neuronale ale unei caracatițe leagă creierul de brațele sale, fiecare dintre acestea conținând milioane de neuroni. După cum a remarcat Godfrey-Smith, nu este nerezonabil să ne gândim la o caracatiță ca având nouă creiere interconectate, câte unul pentru fiecare braț, în plus față de structura cerebrală centrală din cap.

fii acea creatură este acea creatură. Doar Obadiah știe cum este să fii Obadiah. Doar eu știu cum e să fii Arthur Reber - și tu, de asemenea, ai propria ta versiune unică, personală de cunoaștere.

Dar Nagel nu dorea doar să speculeze cu privire la viața mentală a altor specii sau a altor persoane. El avea un scop mai larg, o critică a reducționismului, a materialismului. Nagel susținea că oamenii de știință din domeniul cognitivului, precum subsemnatul, echipați cu modelele reducționiste și mecaniciste cu care ne simțim cel mai confortabil, sunt neputincioși în fața acestui argument „să știm la ce seamănă". Am putea analiza fiecare neuron, fiecare celulă, fiecare bucățică de ADN din Obadiah și tot am fi complet ignoranți cu privire la Obi-ness-ul său esențial. Un neurocercetător surd poate analiza toate aspectele neuroanatomice ale auzului, dar poate fi subiectiv gol atunci când contemplă cum sună numele său. Și mai rău, a argumentat Nagel, suntem la fel de incapabili să oferim o caracterizare obiectivă chiar și a conștiinței noastre personale, umane, în special din cauza relațiilor sale interconectate cu numeroasele operațiuni ale inconștientului care au loc și care au roluri cauzale convingătoare în comportamentele și gândurile noastre. Nu am cum să vă explic vreodată cum este în creierul meu (nu că ați vrea să știți) și este la fel de inutil să încerc să mi-o explic mie însumi! Pe scurt, suntem cu adevărat blocați. Dacă adoptăm o abordare pur subiectivistă, nu vom reuși niciodată să înțelegem nimic altceva decât propriile noastre stări interne. Dacă încercăm să adoptăm o poziție fizicistă și reducționistă, va trebui să lăsăm subiectivitatea pe masă, ca o necunoscută.

Dar, din punctul meu de vedere, cheia este faptul că există ceva asemănător cu a fi un liliac, o caracatiță sau o___(completați spațiul gol după cum doriți). Argumentul pe care îl voi dezvolta este simplu. Deoarece modelul **CBC** adoptă o poziție evoluționistă, interspecifică, care pune sub aceeași umbrelă ontologică sensibilitatea tuturor speciilor, de la cele mai simple la cele mai complexe din punct de vedere cognitiv, enigmele filosofice precum cea cu care ne-a împovărat Nagel sunt neutralizate și, prin urmare, nu sunt relevante pentru argument. Chiar dacă nu putem ști cum este să fii altceva decât noi înșine, chiar dacă suntem limitați la perspectiva noastră științifică, de la persoana a treia, putem totuși să concluzionăm cu ușurință că acest „cum este" este real. Nu este o iluzie. Există ceva ce seamănă cu a fi eu, tu, Obi, unul dintre cefalopodele enervante ale lui Godfrey-Smith care sparg becurile și, desigur, un Paramecium. În analiza finală, asta este ceea ce contează.

Există ceva asemănător cu a fi robot?

Așadar, revenim la întrebarea care ne-a făcut să pornim pe această tangentă: Ce se întâmplă dacă lucrul în cauză nu este o entitate vie bazată pe carbon? Ce se întâmplă dacă nu este organic, ci făcut din altceva, cum ar fi siliciu, roți dințate și scripeți, sfoară și cutii de conserve? Există cum ar fi să fii un robot (foarte sofisticat, complex)? Ar putea fi dezvoltată o inteligență artificială care să aibă o conștiință autentică? Stări subiective interne? Emoții? Contează faptul că unii roboți moderni, articulați, pot efectua acte familiare precum mersul și vorbitul? Capacitatea sa de a răspunde la întrebări care necesită cunoștințe despre lume care se potrivesc cu baza noastră de cunoștințe ne dă de gândit? Dacă cineva ar construi un AI cu toate componentele cauzale ale unei minți umane (sau ale unei pisici), nu ar avea aceleași experiențe umane/feline? Faptul că un computer poate rezolva probleme arată că gândește, înseamnă că are (sau ar putea avea) stări mentale, o minte, conștiință?

Din perspectiva modelului **CBC**, răspunsul la toate aceste întrebări este un „nu" clar, nu există nimic care să semene cu a fi un robot sau orice alt AI. Un AI poate simula gândirea, dar nu îl imită pe gânditor. Există AI-uri care joacă șah și care îi pot învinge pe cei mai buni, dar nu știu nimic despre șah. Dar, de-a lungul anilor, destul de mulți oameni de știință și filosofi au răspuns la aceste întrebări cu răspunsuri care variază de la un „da" entuziast la un „poate" ezitant. Așadar, să aruncăm o privire asupra acestor argumente și a argumentelor împotriva lor.

În primul rând, trebuie apreciat faptul că termenul inteligență artificială și abrevierea sa comună AI sunt utilizate în mai multe moduri diferite. Ocazional, acestea se vor referi la o ramură a epistemologiei care se ocupă cu cunoașterea, dobândirea și reprezentarea acesteia, precum și cu caracteristicile entităților competente din punct de vedere cognitiv instanțiate în artefacte. Aici sunt folosite pentru a se referi la domeniul științific bogat interdisciplinar populat de informaticieni, lingviști, filosofi, antropologi și neuroștiințifici cognitivi care au ca obiectiv principal crearea de sisteme bazate pe calculator care se comportă inteligent. Acest sens se regăsește la un cercetător care, atunci când este întrebat cu ce se ocupă, răspunde: „Lucrez în IA”.

Utilizarea cotidiană mai frecventă este aceea de a face referire la entitatea în sine, cea care efectuează calcule inteligente. Aici se poate face referire la un sistem la fel de simplu ca primele programe care au primit eticheta „AI”, cum ar fi *General Problem Solver* (GPS) dezvoltat în 1959 la Carnegie Mellon de Herb Simon[18], J. C. Shaw și Alan Newell. Funcțiile GPS-ului erau limitate la rezolvarea problemelor logice și geometrice euclidiene și la demonstrarea teoremelor din aceste domenii. Inteligența artificială modernă dispune de un software uimitor de puternic și poate juca jocuri complexe precum șahul și Go, pune diagnostice medicale și recomandă tratamente și chiar „face” știință, analizând baze de date uriașe pentru a descoperi tipare și relații care eludează capacitățile umane și pentru a dezvolta ipoteze care pot fi apoi testate experimental. Ei pot analiza un număr uriaș de puncte de date pentru a obține previziuni meteorologice mai exacte și pot lua decizii bine judecate în situații cu informații parțiale (sau incomplete), cum ar fi războaiele, negocierile de afaceri sau, fascinant, pokerul.

În cele din urmă, termenul va fi utilizat pentru entitatea artificială încă nerealizată care emulează cu adevărat experiența umană, cu abilități de autoreflecție, intenționalitate și capacitate de rezolvare a problemelor pe deplin funcționale - inteligența artificială care va experimenta bucuria, tristețea, dragostea, ura și confuzia; va manifesta o înțelegere reală a abstractului, complexului și frumosului; și își va contempla propria moarte. Pe scurt, inteligența artificială se poate referi la un domeniu științific, precum și la procesele și produsele pe care le generează.

Filosoful de la Berkeley, John Searle, a numit primele exerciții, cele orientate spre scrierea de software pentru a rezolva probleme specifice, *inteligență artificială slabă*, iar ultimele eforturi, cele concentrate pe emularea minții și conștiinței umane, *inteligență artificială puternică*. Toate programele de calculator existente în prezent sunt exemple de inteligență artificială slabă, indiferent cât de sofisticate sunt. Printre progresele recente se numără *Libratus* de la Carnegie Mellon, un robot care joacă heads-up, *no-limit poker*[19] mai bine decât cei mai pricepuți jucători umani (pentru detalii, a se vedea Brown & Sandholm, 2018), AI-ul *DeepStack* al *Computer Poker Research Group* de la Universitatea din Alberta, care a învins 10 din 11 jucători profesioniști de rang înalt (a fost, de asemenea, în fața ultimului, dar nu suficient de mult pentru a fi considerat statistic fiabil; pentru detalii, a se vedea Moravcik et al. 2017) și *AlphaGo* de la Google, care a cucerit recent campionatul mondial la Go. Acestea au fost precedate, în mod celebru, de *Deep Blue*, inteligența artificială IBM

18 Simon a fost un adevărat polimat și un contribuitor extraordinar de influent în mai multe domenii, aparent disparate. Pe lângă contribuțiile fundamentale aduse informaticii, științelor cognitive și psihologiei organizaționale, Simon a primit Premiul Nobel pentru Științe Economice în 1978. În mod amuzant, doctoratul său, de la Universitatea din Chicago, nu a fost în niciunul dintre aceste domenii (ci în științe politice).

19 O formă de poker în care există doar doi jucători în fiecare mână, iar singura limită privind suma care poate fi pariată în orice moment este numărul de jetoane pe care îl aveți. Este un joc diabolic de complex. Spre deosebire de jocuri precum șahul și Go, pokerul este un joc cu informații parțiale sau incomplete, ceea ce face deosebit de dificilă dezvoltarea unei inteligențe artificiale care să îl joace cu succes. La șah și go vezi tabla și știi unde sunt toate piesele adversarului. La poker, nu cunoști cărțile adversarului și trebuie să faci deducții pe baza unor elemente precum modelele de pariere și acțiunile anterioare, inclusiv cele care reflectă stilul personal. Până în prezent, roboții de poker de succes joacă doar „heads-up”, împotriva unui singur adversar. Dacă se adaugă un al treilea jucător, jocul devine prea complex pentru programele AI existente. Cel mai obișnuit joc de poker are nouă jucători și, prin urmare, roboții de poker sunt departe de a putea concura la mese complete.

care joacă șah și care, în 1997, l-a învins pe Garry Kasparov și Watson, inteligența artificială care i-a învins cu ușurință pe cei mai buni jucători de *Jeopardy* pe care i-a produs omenirea.

Este posibil să fi observat că oamenii de știință din domeniul informaticii și al roboticii se simt irezistibil tentați să dea nume AI-urilor lor. Acest lucru este de înțeles într-un sens, deoarece denumirea ne oferă o modalitate rapidă de a vorbi despre AI fără referințe lungi precum „AI-ul care joacă Go” sau „AI-ul care l-a învins pe Kasparov”. Dar bănuiesc că este vorba despre o psihologie subtilă. Se pare că atribuirea unui „nume” unui AI are legătură cu tendințele noastre de a antropomorfiza și de a conferi obiectelor ceea ce se numește „agenție”. Vom analiza mai detaliat aceste aspecte în continuare. Aceasta ne va ajuta să înțelegem ușurința cu care ne putem imagina o IA conștientă.

Aceste sisteme AI de joc nu au fost dezvoltate doar pentru a juca jocuri. Ele au o generalitate considerabilă. Algoritmii încorporați în jocul de poker *Libratus*, de exemplu, sunt aplicați în prezent în alte domenii cu informații incomplete sau parțiale, cum ar fi diagnosticul și tratamentul medical, deciziile strategice de apărare, evaluările ample ale piețelor economice care implică evaluarea riscurilor și proiectarea tendințelor financiare. Acestea sunt unele dintre cele mai dificile probleme pentru inteligența artificială, deoarece doar o parte din informațiile necesare pentru a lua decizii optime sunt cunoscute și fiecare dintre participanți știe lucruri diferite. În poker, ca și în război, nu știi ce cărți sau arme deține adversarul tău, la ce s-ar putea gândi sau ce ar putea gândi despre ceea ce gândești tu. În diagnosticul medical, o mare parte din trecutul pacientului nu este disponibilă, doar unele dintre simptomele posibile pot fi prezente și acestea prezintă variații considerabile de la pacient la pacient. În prognozarea piețelor, variabilitatea naturală din cadrul sistemului este atât de mare încât nu poate fi complet instanțiată în program. În schimb, programul utilizează rutine care scanează și captează covariațiile statistice dintre elementele din datele de intrare. Unele utilizează euristici precum cele folosite de oameni, altele utilizează proceduri foarte diferite, dar cu câștiguri enorme în viteză și putere de calcul. Aceste domenii sunt, de asemenea, printre cele mai greu de stăpânit de către oameni.

În plus față de aceste realizări, lucrările recente privind ceea ce a fost numit „învățarea profundă” utilizează algoritmi puternici care se bazează pe modele și caracteristici structurale ale mediilor complexe pentru a extrage chiar și cele mai timide legături stohastice între aspectele domeniului căutat. Mașinile de învățare profundă au fost utilizate în cercetare în domenii la fel de diverse precum prelucrarea limbajului, traducerea automată și viziunea pe calculator. De asemenea, acestea au fost utilizate, printre multe alte aplicații, pentru a căuta noi medicamente, pentru a determina care dintre un grup de terapii medicamentoase este probabil să fie eficientă pentru un anumit pacient, pentru a oferi soluții la probleme interpersonale și organizaționale complexe și pentru a face evaluări și previziuni meteorologice. Cele mai multe dintre acestea se bazează pe arhitecturi de rețele neuronale *multistrat* și sunt adesea superioare oamenilor în ceea ce privește detectarea modelelor și a relațiilor în contexte complexe, multivariabile. Dar toate aceste progrese, precum și cele care vor urma, sunt rezultatul puterii brute de măcinare a unor șiruri foarte lungi de zerouri și unu. Aici nu există sensibilitate. Nici o înțelegere identificabilă epistemic. Avem de-a face cu forța brută a unei mașini lipsite de orice sentiment de sine sau de conștientizarea a ceea ce se realizează.

În mod interesant, Michael Bowling, directorul echipei de cercetare de la Universitatea din Alberta, Edmonton, s-a referit la modul în care *DeepStack* ia decizii în jocul său de poker ca fiind „intuitiv”. În loc să scaneze toate elementele posibile ale jocului de poker, care sunt aproape nenumărabil de mari, acesta evaluează fiecare situație pe care a întâlnit-o ca și cum ar fi un minigame. La fel cum un dermatolog se uită la o erupție cutanată și emite o judecată „intuitivă” în legătură cu aceasta, pe baza unui grup de pacienți similari văzuți în trecut cu tipuri similare de erupții sau leziuni cutanate, *DeepStack* se uită la o mână de poker și decide cum să o joace, pe baza unui grup de situații similare jucate împotriva acestui adversar în trecut. În mod impresionant,

DeepStack (ca orice medic bun) își revizuiește deciziile de poker corecte (sau diagnosticele corecte) și reevaluează mâinile jucate greșit (sau diagnosticele greșite) și se ajustează astfel încât, în timp, precizia să crească și rata de eroare să scadă.

Dar, în ciuda faptului că am folosit cuvinte precum „a privi” și „a decide”, iar Bowling a scris că AI-ul grupului său de cercetare „ia decizii” bazate pe „intuiție”, toate acestea sunt exemple de AI slabe. *Libratus* nu știe nimic despre impactul emoțional al unei mari cacealmale în poker; de fapt, nu știe nimic despre poker. *Deep Blue* nu are o înțelegere reală din punct de vedere epistemic a ceea ce este un final de joc în șah. Un program de diagnosticare ar putea furniza informații medicale utile, dar face acest lucru fără să aibă habar de conceptul de boală. Pentru a împrumuta o distincție utilă a lui Daniel Dennett, există competență, dar nu și înțelegere.

Inteligența artificială slabă se referă, în cea mai mare parte, la program, nu la dispozitivul pe care este rulat sau la reprezentările sale bazate pe cunoștințe - deși există circumstanțe în care dispozitivul este important. Uneori, hardware-ul va fi dedicat rutinelor din program (ca în cazul *Deep Blue* sau *Watson*) sau sunt necesare bănci mari de memorie pentru această sarcină (*Libratus*, predicții meteorologice), dar acestea sunt convenienţe practice, nu necesități teoretice. Computerul pe care rulează un program este un dispozitiv util pentru că efectuează rapid calcule complexe. În principiu, victoria lui *Deep Blue* asupra lui Kasparov ar fi putut fi realizată cu creion, hârtie și un abac. Ar fi durat foarte mult timp, dar rezultatul ar fi fost același. De fapt, atunci când Turing a dezvoltat pentru prima dată teoria unei „mașini de calcul universale”, aceasta a fost descrisă ca o bandă fără sfârșit, împărțită în unități, care se deplasa sub un cap de citire/scriere în funcție de instrucțiunile privind deplasarea la stânga sau la dreapta. Informaticianul Joseph Weizenbaum a observat că o mașină Turing ar putea fi instanțiată de o rolă foarte lungă de hârtie igienică și câteva pietricele. *Inteligența artificială slabă* se referă la modele de comportament inteligent. *IA puternică* se referă la entități artificiale cu anumite stări mentale ontologice. Întrebarea-cheie este dacă o entitate din prima varietate ar putea deveni una din a doua dacă este programată în modul potrivit. Filosoful John Searle de la Berkeley a prezentat un argument inteligent în urmă cu aproape 40 de ani, pe care va fi util să îl revizităm.

Camera chinezească a lui John Searle

Cel mai simplu mod de a înțelege această distincție între *inteligență artificială slabă* și *puternică* este să ne întoarcem la experimentul *Gedanken* (sau „gândire”) introdus de Searle în 1980: *Camera chineză*. Searle ne-a invitat să ne imaginăm pe cineva, să-i spunem Arnie, stând într-o cameră închisă (a se vedea figura 1.1).

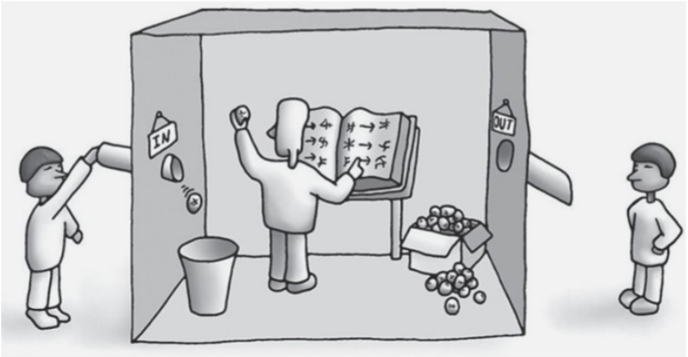

Figura 1.1 Camera chineză a lui Searle. Un vorbitor de chineză trimite întrebări. Arnie își consultă manualul de instrucțiuni și, în cele din urmă, va emite un răspuns în ceea ce un alt vorbitor de chineză va vedea că este o chineză perfectă. Wikicommons, open-source. Desen original realizat de American Pink Images (https://america.pink/images/9/6/3/2/5/4/en/2-chinese-room.jpg).

Știm câteva lucruri despre el, unul dintre acestea fiind că este un vorbitor monolingv de limba engleză care nu are absolut nicio cunoștință de chineză, nu o poate citi, scrie sau vorbi. Arnie nu are niciun contact cu lumea exterioară, cu excepția foilor de hârtie care sunt schimbate printr-o fantă în unul dintre pereții camerei sale. De asemenea, are un set de instrucțiuni despre cum să manipuleze simbolurile pentru a răspunde la orice mesaj care intră prin fantă. Curând, încep să apară bucăți de hârtie. Fără ca Arnie să știe, fiecare este o întrebare scrisă în chineză. El își consultă cărțile de instrucțiuni, care îi spun ce fel de simboluri să compună pentru fiecare intrare în parte. Ho le notează conștiincios, din nou fără să știe ce a făcut și strecoară răspunsurile.

Să presupunem acum că setul de instrucțiuni de manipulare a simbolurilor cu care lucrează Arnie este atât de bine construit încât fiecare întrebare pusă primește un răspuns adecvat în limba chineză gramaticală. Pentru cineva care observă secvența de întrebări și răspunsuri, ar fi normal să concluzioneze că Arnie vorbește fluent chineza. Dar acest lucru este pur și simplu greșit. Am pornit de la premisa că este un vorbitor monolingv de limba engleză. El nu răspunde cu adevărat la întrebări în chineză; el doar manipulează simboluri, urmând un set de linii directoare de calcul sintactic (deși nici el nu știe acest lucru), care se întâmplă să satisfacă constrângerile lingvistice ale limbii chineze. Tot ceea ce face este să se uite la un set de mâzgăleli fără sens (pentru el) și să așeze pe o bucată de hârtie un alt set de mâzgăleli la fel de lipsite de sens (pentru el). Nu există nicio înțelegere aici, nicio semantică. După cum spunea filosoful Owen Flanagan, „nu există nimic care să semene cu a fi Camera Chineză”.

Rețineți, nu contează cine sau ce se află în Camera Chineză. Am avut un om numit Arnie, dar ar fi putut fi Searle însuși. Ar fi putut fi un computer. De fapt, nu puțini dintre criticii lui Searle au subliniat că experimentul de gândire original nu are prea mult sens din perspectiva lumii reale. Arnie ar trebui să aibă o bază de date imensă cu zeci, sute de mii, poate milioane de foi de hârtie stocate în cine știe câte dosare și dulapuri de arhivare și că i-ar lua ore, zile... săptămâni pentru a răspunde la orice întrebare moderat de interesantă. Acest lucru este adevărat, dar irelevant; acesta este, la urma urmei, un experiment de gândire conceput pentru a acționa ca, în expresia utilă a lui Daniel Dennett, o „pompă de intuiție”. Putem pune orice entitate dorim în cameră (alta decât un vorbitor de limba chineză alfabetizat) și două lucruri ar fi în continuare clare: (a) ce/cine ar fi ar trece testul Turing deoarece observatorii externi ar concluziona că înăuntru se află un vorbitor de limba chineză autentic, competent din punct de vedere lingvistic și (b) tot nu ar exista nicio „cunoaștere” acolo - doar o secvență de operațiuni în care simbolurile sunt deplasate, urmând reguli sintactice. Pe scurt, deși camera chineză se comportă ca un AI cu experiențe cognitive autentice (și anume, ca o instanță a unui AI puternic), nu este. Efectuarea calculelor nu este suficientă.

Camera chineză a lui Searle a fost supusă unor critici viguroase din diverse părți. Ea a fost analizată de filosofi, neuroștiințifici, informaticieni, antropologi și psihologi. În 2002, filosoful David Cole a afirmat că „argumentul camerei chinezești a fost probabil cel mai discutat argument filosofic din domeniul științelor cognitive care a apărut în ultimii 25 de ani”. Comentariul lui Cole a apărut într-un volum dedicat în întregime discuțiilor (în mare parte critice) despre *Chinese Room* (a se vedea Preston & Bishop, 2002).

Desigur, în virtutea premiselor de bază ale modelului **CBC**, m-am trezit membru al unei mici alianțe, a celor cu înclinații filosofice care au fost de acord din start cu ipoteza centrală a lui Searle. Există, de asemenea, o mică notă personală relevantă aici. După cum am menționat mai devreme, unul dintre participanții la Institutul Național pentru Științe Umaniste din 1981 privind psihologia și filosofia minții a fost Searle, care, desigur, ne-a oferit argumentele sale nou inventate, care au apărut din cadrul *camerei chinezești*, împotriva posibilității de a instanția o IA puternică adevărată. Deși a fost imediat atacat de aproape toți cei cu pregătire filosofică, a fost acceptat în mod la fel de natural ca argument convingător de către, dacă memoria nu mă înșeală, toți cei cu experiență în psihologia cognitivă. Am găsit această diviziune fascinantă atunci și o găsesc și astăzi. După cum s-a menționat în altă parte, filosofii sunt instruiți în dezbateri și se pare că le practică când și unde pot. Și, după cum am menționat în altă parte, cred că acesta este un lucru foarte bun.

Dar a fi de acord cu Searle atunci (și încă și astăzi) nu înseamnă a încheia afacerea. După cum vom afla din acea discuție cu omida, el nu a mers suficient de departe. Searle este/era mulțumit

să afirme că „creierul face mintea”. Așa este. Dar ele nu sunt singurele care fac acest lucru și, limitându-și explorările la creiere, în special la creierele umane, Searle a dansat în jurul lor, dar nu a intrat niciodată în domeniul mai profund care mă interesează pe mine.

Este rezonabil să vă întrebați de ce v-am târât pe această cale. De ce această șmecherie filozofică inteligentă, veche acum de peste 35 de ani, este o critică serioasă a inteligenței artificiale puternice? Motivul este că *inteligența artificială puternică* este o propunere independentă de hardware. Ea este îmbrățișată de o școală epistemologică care poartă mai multe nume: funcționalism, funcționalism independent de hardware, funcționalism informatic, computaționalism sau teoria computațională. Principala axiomă aici, cea la care Putnam a făcut aluzie în urmă cu peste o jumătate de secol, este că nu contează ce dispozitiv efectuează calculele sau îndeplinește funcțiile. Douglas Hofstadter, venind din partea cercetătorilor cognitivi ai funcționalismului, a spus-o succint (a se vedea Hofstadter. 2007 și Hofstadter & Dennett, 1981 pentru mai multe informații despre opiniile lui Hofstadter):

Mințile există în creiere și pot ajunge să existe în mașini programate. Dacă și când astfel de mașini vor apărea, puterea lor cauzală nu va proveni din substanțele din care sunt făcute, ci din designul lor și din programele care rulează în ele.

Christoph Koch a intrat în luptă din postura de neurocercetător:

Indiferent dacă organismul sau artefactul provine din regnul antic Animalia sau din descendența sa recentă din siliciu, indiferent dacă are picioare pentru a merge, aripi pentru a zbura sau roți cu care să se rostogolească - dacă are atât stări diferențiate, cât și stări integrate ale informației, simte ceva ca fiind un astfel de sistem; are o perspectivă interioară.

La fel ca David Chalmers - venind din perspectiva filosofului minții:

Nu contează dacă organizarea [sau entitatea] este realizată în cipuri de siliciu, în populația Chinei sau în cutii de bere și mingi de ping-pong. Atât timp cât organizarea funcțională este corectă, experiența conștientă va fi determinată.

Din acest punct de vedere, calculatoarele ar trebui cu siguranță să fie capabile să devină conștiente și să aibă stări mentale. La fel și entitățile cristaline sau rețelele de corzi și cutii de conserve, atâta timp cât efectuează calculele corecte, emulează rețelele neuronale corecte cu cauzele lor comportamentale corespunzătoare și instanțează reprezentările corecte. Argumentul lui Searle este că nu se poate obține semantica din sintaxă. Hofstadter și Chalmers spun că se poate - dacă se efectuează operațiile sintactice corecte. Dar, cu toată agitația, bănuiesc că valul de critici a fost în principal un exercițiu de exercitare. Filosofilor le place să se certe. Argumentele ne mențin cinstiți, iar adevărata sarcină a filosofiei este să mențină pe toată lumea (inclusiv pe ei înșiși) cinstită intelectual. Propunerile trebuie să fie contestate și puse la încercare, pentru a vedea ce supraviețuiește și ce se ofilește și se ascunde în întuneric.

Pe măsură ce praful se așează încet (după cum s-a menționat, această dezbatere durează de ceva timp), ideile de bază ale lui Searle sunt, în cele mai multe cercuri, acceptate ca fiind în esență corecte - chiar și Putnam, cel care a început toată această poveste, a ajuns să fie de acord. Da, creierele creează minți. Ele fac acest lucru nu din cauza a ceea ce fac, ci pentru că sunt alcătuite din materialul potrivit pentru crearea minții și sunt organizate în modul potrivit. Sistemul este extrem de dependent de hardware (sau, dacă preferați, de *netware*). De la cele mai elementare experiențe reactive ale unui organism unicelular care reacționează la mediul înconjurător până la formele sofisticate de autoreflecție care caracterizează specia noastră, totul depinde de existența materiei

organice potrivite, structurată și organizată în modul potrivit. Trebuie să ai materia potrivită pentru a obține experiență, doar amestecul corect de molecule pentru sensibilitate, doar echilibrul precis de proteine și enzime pentru subiectivitate. O inteligență artificială bazată pe siliciu poate acționa inteligent, dar este proastă ca o cărămidă. Și probabil destinată să rămână așa.

Există, desigur, persoane care încă susțin că o anumită formă de inteligență artificială este, în principiu, capabilă să devină conștientă, dezvoltând o sensibilitate ontologică - printre aceștia se numără unii gânditori proeminenți, inclusiv psihologul cognitiv Stanislaw Dehaene, filosoful Daniel Dennett și neurocercetătorii Giulio Tononi și, după cum am menționat mai sus, Christof Koch.

Dehaene și colegii Hakwan Lau și Sid Kouider au susținut recent că trebuie să facem distincția între o varietate de conștiință care selectează „informații pentru difuzare globală” și una care „monitorizează singură” aceste operațiuni. Ei susțin că inteligența artificială actuală nu este capabilă să simuleze niciuna dintre aceste stări fenomenale și au avut succes doar cu calcule implicite sau inconștiente. Cu toate acestea, ei nu sunt convinși că mașinile conștiente nu sunt posibile, ci doar că generația actuală nu utilizează probabil arhitectura mașinii corespunzătoare. Dennett adoptă o poziție similară, susținând că motivul pentru care sistemele artificiale actuale nu au atins subiectivitatea este acela că accentul în domeniu a fost pus pe puterea brută de calcul. În opinia sa, aceasta a fost o abordare greșită și nu va funcționa „*... deoarece ceea ce ... [dispozitivele actuale de inteligență artificială] sunt alcătuite este organizat în ierarhii greșite*”.

Atât Dehaene, cât și Dennett inserează, în esență, marcaje conceptuale sugerând că un artefact conștient nu ar trebui exclus doar pentru că nu am produs încă unul. Sunt de acord că una dintre cheile apariției sensibilității se va găsi probabil în modul în care sunt organizate anumite elemente și caracteristici ale unui sistem, dar încă nu am văzut niciun argument convingător sau set de date care să sugereze că astfel de funcții producătoare de sensibilitate pot avea loc în entități inerte, artificiale. După cum am menționat mai sus și după cum vom susține în restul volumului, chestiunea în sine este esențială. Organizarea este esențială, dar, pentru a vedea apariția unei subiectivități veritabile, aceasta trebuie să fie alcătuită din lucrurile potrivite, din materialele potrivite, iar acestea vor fi **molecule organice**.

Tononi și Koch urmează o cale similară, în sensul că se concentrează mai degrabă pe structură decât pe funcțiile pur computaționale. Poziția lor se bazează pe un model dezvoltat de Tononi, cunoscut sub numele de teoria informației integrative (**IIT**), care se bazează pe ideea că sensibilitatea și funcționarea cerebrală se bazează pe informație și că, dacă sistemul, orice sistem, este capabil să gestioneze suficientă informație, acesta va manifesta sensibilitate, cu condiția ca părțile componente să fie interactive sau, după cum se spune în numele teoriei, „integrate”. Vom analiza mai îndeaproape **IIT** în capitolul 2, deoarece a avut un impact asupra modului în care unii biologi și neuroștiințifici abordează problema dificilă. O vom găsi deficitară din mai multe motive, de la probleme legate de validitatea axiomelor modelului până la faptul că, prin însăși natura sa, i-a determinat pe Tononi și Koch să îmbrățișeze panpsihismul. Știu, această propoziție este criptică. Ideea va fi clarificată în capitolul 2.

Înainte de a trece mai departe, permiteți-mi să mai fac o scurtă plimbare pe o cale ușor diferită - una care apare din perspectiva sociologului. În viitoarea sa carte *Artifictional Intelligence: Against Humanity's Surrender to Computers*, sociologul științei Harry Collins susține că, deoarece AI-urile sunt entități informatice izolate, soiurile de inteligență pe care le manifestă, indiferent cât de puternice devin din punct de vedere computațional, vor fi întotdeauna fundamental diferite de inteligența umană deoarece nu au aspecte sociale interactive - de aici neologismul amuzant din titlu. Specialitatea lui Collins este cunoașterea socială, în sensul de cunoaștere care apare într-un context social, cunoaștere care nu ar putea avea loc independent de elementul social. Lucrările sale

anterioare au explorat componentele sociale ale cercetării privind undele gravitaționale, examinând modul în care oamenii de știință care lucrau la o problemă complexă din punct de vedere teoretic, metodologic și empiric interacționau și, în mod colectiv, obțineau informații, cunoștințe și înțelegere (a se vedea Collins, 2004). Mai recent, Collins a analizat modul în care aceste tipuri de cunoștințe sunt dobândite, în mod fascinant, prin interacțiuni intense între oamenii de știință, care adesea nu știu ce au învățat, de fapt. În *Tacit and Explicit Knowledge*, Collins a dezvoltat aceste idei într-un mod care se potrivește cu cercetările mele privind învățarea implicită, în care cunoștințele sunt acumulate independent atât de procesul, cât și de produsele învățării.[20]

Punctul de vedere al lui Collins cu privire la IA, deși provine dintr-un domeniu academic destul de diferit, se potrivește perfect cu cel al modelului **CBC** și întărește argumentele prezentate aici. Poziția lui Collins cu privire la inteligența artificială este destul de directă, citând din viitoarea carte:

> *a) Niciun computer nu va vorbi fluent un limbaj natural, nu va trece un test Turing sever și nu va avea o inteligență umană deplină decât dacă va fi complet integrat în societatea umană normală.*
> *b) Niciun computer nu va fi complet integrat în societatea umană ca urmare a progresului incremental bazat pe tehnicile actuale*.

Rețineți că Collins nu este interesat de subiectele alese de mine, conștiința și mintea. El se concentrează pe inteligență, cunoaștere, modul în care se dobândește expertiza, cât acces la aceste reprezentări epistemice au cunoscătorii și modul în care acestea depind de un context social și sunt încorporate inextricabil în acesta. Dar acest lucru este în regulă. El este un sociolog al științei, nu un om de știință cognitivă și nu este preocupat în mod special de chestiuni precum filosofia minții (aceste focare diferite au fost, desigur, parțial responsabile de neînțelegerea noastră inițială).

Dar punctul său de vedere se potrivește cu cel al **CBC** prin faptul că adaugă un motiv diferit, dar compatibil, pentru a se îndoi de posibilitatea unei IA puternice sau de viabilitatea perspectivei independente de hardware.

Pe scurt, se pare că este puțin probabil ca obiectivul *inteligenței artificiale puternice* să fie atins, cu siguranță nu prea curând - dacă nu chiar niciodată. Dar să ne asigurăm că am lămurit un aspect. Deși susțin argumentul că obiectivul *inteligenței artificiale puternice*, Sfântul Graal al funcționaliștilor, este probabil, în principiu, de neatins, nu spun că un sistem mecanicist, bazat pe fizică, nu poate fi conștient. Noi, tu și cu mine, suntem doar astfel de entități. Suntem entități fizice, fiziologice, biologice. Suntem supuși forțelor și proceselor biologiei evolutive. Împărtășim un patrimoniu bio-genetic cu toate celelalte specii de animale de pe planetă. Cu toții suntem simțitori. Argumentul împotriva *IA puternică* nu subminează în niciun fel o poziție obiectivă, fizicistă, reducționistă. Noi și toate celelalte organisme cu diferite grade de agilitate mentală suntem entități biologice, iar capacitățile noastre afective, comportamentale și cognitive nu sunt nici mai mult (nici mai puțin) decât suma operațiunilor și interacțiunilor lor complexe. Pot mașinile să gândească? Pot fi inteligente? Pot fi conștiente? Bineînțeles că pot; noi și orice altă specie vie suntem doar astfel de mașini, dar construite pe o platformă organică adecvată.

Interesant este faptul că această farsă intelectuală pe care filosofii și unii neuroștiințifici au găsit-o irezistibilă nu a jucat practic niciun rol în rândul celor care lucrează efectiv în domeniul inteligenței artificiale. De la începutul anilor 1980, informaticienii din cadrul acestei specialități au

---

20 Interesant este faptul că, în ciuda suprapunerii evidente dintre cercetările lui Collins și ale mele privind dobândirea de cunoștințe tacite, ne-am aflat în conflict atunci când am încercat să ne comunicăm reciproc perspectivele. Această ciocnire a punctelor de vedere, care se apropia de un „moment incomensurabil" kuhnian, ne-a determinat să încercăm să deslușim natura discursului interdisciplinar. A se vedea Collins & Reber (2013) pentru modul în care am reușit să facem acest lucru, am ajuns să ne înțelegem reciproc modul de a folosi limbajul și de a face știință și, pe parcurs, am devenit buni prieteni.

fost mai preocupați să afle dacă programele lor fac ceea ce vor să facă, dacă acționează inteligent, dacă iau decizii utile, dacă găsesc soluții logice și dacă învață eficient despre datele de intrare la care sunt supuse. Programele au devenit din ce în ce mai puternice, unele foarte specializate, altele scrise cu scopuri generale și multifuncționale în minte, însă cercetătorii care le dezvoltă se concentrează aproape în întregime asupra modalităților de realizare a obiectivelor identificabile, independent de modul în care oamenii ar putea să le atingă. Puțini și-au făcut griji cu privire la problemele epistemice care i-au preocupat atât de mult pe filosofi. Ediția din 7 iulie 2017 a revistei *Science* a publicat pe copertă articolul „*AI transforms science*”. Ediția a inclus un editorial, nouă articole despre progresele recente și o duzină de scrisori de la cercetători care își rezumă activitatea. Nimeni nu a menționat conștiința.

Interesant, chiar și cercetătorii care se concentrează pe cercetarea de vârf în domeniul inteligenței generale artificiale (AGI) vorbesc rar despre conștiință. O **AGI** ar fi una care funcționează ca un om. Ar fi capabilă să învețe despre circumstanțe și situații noi odată ce a primit educația corespunzătoare, instruirea relevantă. Acestea nu ar fi AI slabe în terminologia lui Searle, deoarece nu ar fi legate de anumite sarcini. Dar, în mod curios, nu se discută în mod obișnuit dacă acestea ar fi sau nu puternice. Cercetătorii în domeniul **AGI** tind să se concentreze asupra eforturilor de a face din dispozitivele lor entități generaliste, competente și neutre din punct de vedere al domeniului. Dacă un **AGI** complet funcțional în viitor ar fi conștient în același sens ca și oamenii pe care i-a simulat nu face în mod normal parte din discuție - cu excepția cazului în care, desigur, sunt prezenți filosofi ai minții.

Un motiv întemeiat pentru această poziție în cadrul comunității IA mai largi este un sentiment prezentat de bio-computaționalistul James Reggia. Reggia face parte dintr-o cohortă din ce în ce mai numeroasă de cercetători în domeniul IA care lucrează la dezvoltarea de tehnici de calcul pentru simularea operațiunilor biologice naturale, cum ar fi rețelele neuronale suspectate a fi unitățile funcționale ale operațiunilor cerebrale, mecanismele evolutive care conduc procesul de speciare și mecanismele care acționează în reacțiile imunitare. Unele dintre aceste cercetări devin foarte ezoterice; diverse proiecte încearcă să capteze în limbaj computațional mecanismele care controlează comportamentul de roi la lăcuste, cele care permit coordonarea optimizată în coloniile de furnici și posibilitatea instanțierii vieții artificiale și a autoreproducerii. Într-o prezentare generală a domeniului din 2013, Reggia a recunoscut că s-au făcut progrese considerabile în dezvoltarea unor modele eficiente de *inteligență artificială slabă*, care pot servi drept platforme pentru studiul conștiinței și sunt capabile să ofere simulări automate ale proceselor neurocognitive implicate în percepție, învățare și memorie. Însă, în ceea ce privește pretențiile funcționaliste ale sensibilității independente de hardware, acesta a concluzionat:

[Nicio abordare existentă a conștiinței artificiale nu a prezentat o demonstrație convingătoare a conștiinței fenomenale a mașinii sau chiar dovezi clare că conștiința fenomenală artificială va fi posibilă în cele din urmă.

„Tragerea” antropomorfă

De ce au îmbrățișat atât de mulți funcționalismul computațional și de ce există încă persoane reticente în a renunța la această propunere? În termenii cei mai simpli, bănuiesc că este din cauză că există o atracție emoțională, antropomorfică, o atracție afectivă care ne face să dorim să dotăm roboții cu conștiință, conștientizare - și există motive psihologice întemeiate pentru aceasta.

Această atracție există și este reală. Oamenii obișnuiți o simt și adesea se bucură de ea. Oamenii de știință și filosofii o simt și ei (eu cu siguranță nu sunt imun) - deși, bănuiesc, mulți încearcă să îi atenueze impactul sau au învățat să îi neutralizeze atracția. Este una dintre principalele atracții ale unui gen din *science fiction*, romane și filme futuriste și povești fantastice în care mințile

și corpurile umane se contopesc sau se amestecă cu roboții pentru a produce hibrizi om-mașină. Este, de asemenea, o provocare pentru scriitorii care descriu viitoruri apocaliptice în care inteligența artificială sensibilă, folosind nanotehnologia, genetica manipulatoare și alte tehnici nebănuite, devine atât de inteligentă și de deșteaptă decât oamenii, încât devine stăpânul nostru și amenință să preia controlul asupra planetei - un scenariu obișnuit pentru filmele „*blockbuster* de vară".

Informaticianul, inovatorul, scriitorul și futuristul Ray Kurzweil este unul dintre cei mai cunoscuți susținători. Cartea sa *The Singularity is Near* (Singularitatea este aproape) a fost un succes (și filmul cu același nume) în care inteligența artificială devine mult mai puternică decât oamenii și își extinde raza de acțiune, iradiind în spațiu până când universul devine o entitate pe deplin conștientă. În cartea „*Mind Children*" a informaticianului Hans Moravec, am încărcat minți umane în programe pentru roboți, iar în cartea sa „*Robot: Simple Machine to Transcendent Mind* (De la simplă mașină la minte transcendentă) prezintă roboți superinteligenți care sculptează o lume în care sensibilitatea pură înlocuiește în cele din urmă formele fizice. În timp ce aceste previziuni neîngrădite ale vieții în viitor sunt captivante, ele pot, de asemenea, după cum a remarcat filosoful Colin McGinn într-o recenzie a cărții *Robot* a lui Moravec, „să se transforme maiestuos în incoerență".

Și, bineînțeles, atracția se simte chiar și în cazurile în care scenariștii recunosc ce se întâmplă și se luptă cu asta. Personajul Data din *Star Trek: The Next Generation* este cu siguranță cel mai bun (și cel mai cunoscut) exemplu. Interpretat de Brent Spiner, Data era un android conștient de sine, un robot umanoid. Făcându-l pe Data să pară pe deplin uman, tendința de a acorda stări mentale umane entităților care arată ca noi a fost stimulată. Dar, pentru a-l descrie pe Data ca nefiind pe deplin uman, personajul a fost creat și fără emoții.[21]. Această separare a stărilor interne afective de cele cognitive se încadrează perfect în orientarea computațională a funcționaliștilor.

Poate că nu ați observat, dar toate discuțiile anterioare cu privire la posibilitatea ca o entitate să aibă o conștiință autentică, independentă de componența sa materială, se refereau la faptul dacă inteligența artificială gândește, cunoaște, calculează, învață, stochează și recuperează date, toate acestea fiind tratate ca forme complexe de procesare a informațiilor. Rareori a fost abordată problema emoțiilor, afectelor, sentimentelor. A fost ca și cum stările emoționale se aflau în afara modelelor computaționaliste[22]. Gândirea poate fi reprezentată doar ca o mulțime de evenimente digitizate rapide în care neuronii (sau tranzistorii) stimulează sau inhibă alți neuroni (sau tranzistori), însă viața emoțională pare a fi mult mai analogică. Simțim puțină durere sau multă, simțim o mică atingere de simpatie sau ne îndrăgostim iremediabil, avem o zi deprimantă sau ne trezim cuprinși de o depresie cruntă. Funcționaliștii s-au ocupat de aceste probleme în mare parte ignorându-le. Dar scenariștii din *Star Trek: The Next Generation* nu au făcut-o. Ei au luat computaționalismul drept ghid și l-au făcut pe Data inteligent, cu abilități cognitive mai mult decât umane, dar lipsit de emoții umane.[23]

Aceste mici manevre teatrale nu sunt accidente. Ele au la bază mecanisme vechi din punct de vedere evolutiv care depind de procesele de recunoaștere, mecanisme care joacă, de asemenea, un rol în tendința de a introduce caracteristici umane în entități artificiale. O privire rapidă asupra a două dintre acestea poate pune în lumină ceea ce se întâmplă aici. Unul este recunoașterea

---

21 Până când un „cip emoțional" a fost instalat într-un episod târziu. Dar, în ciuda incapacității evidente a primului Data de a forma o relație amoroasă, personajul a devenit un sex simbol. După cum a descoperit reporterul Stephanie Dubois, cea mai mare parte a corespondenței lui Spiner/Data provenea de la femei și multe dintre ele erau de natură vădit romantică.

22 De asemenea, a lipsit cu desăvârșire din dezbateri faptul că experiența subiectivă nu este ceva ce are loc doar în interiorul unui creier. Creierele există în corpuri și, după cum a remarcat Evan Thompson, o imagine completă a conștiinței necesită o reprezentare întrupată. În *The Primacy of Movement,* Maxine Sheets-Johnstone a făcut o afirmație similară, subliniind că creierul nu este static. Ele sunt în corpuri, iar corpurile se mișcă și se schimbă. Vom avea mai multe de spus despre aceste aspecte ale conștiinței mai târziu. Încarnarea și mișcarea vor fi elemente importante în modelul **CBC**.

23 Cu toate că presiunea intrigii a dus adesea la derapaje și l-a făcut pe Data să treacă peste această linie. El s-a arătat preocupat de camarazii săi, a fost loial până la un punct și a demonstrat o capacitate muzicală considerabilă, cu elemente emoționale clare - trăsături și aptitudini care sunt greu de imaginat a fi instanțiate în orice entitate lipsită de emoții.

conspecificilor (membri ai aceleiași specii), iar celălalt este detectarea agenției. Deși nu există multe cercetări care să stabilească o legătură directă între aceste mecanisme și proiecția mentalității în roboți, ceea ce există este cu siguranță de susținere.

Recunoașterea conspecifică.

După cum am menționat mai devreme, practic fiecare specie are un sistem de procesare a recunoașterii care este adaptat la conspecifice și aceasta din motive darwiniene întemeiate. Ca factor al strategiei de împerechere a unei specii, acesta reduce probabilitatea unor eforturi ineficiente de reproducere. Căutarea unui partener care să aibă același genom maximizează probabilitatea de a avea urmași viabili și fertili. Ca element al unei strategii sociale, permite indivizilor să identifice alte persoane care sunt mai puțin susceptibile de a fi prădători sau invadatori și, prin urmare, reduce poziția defensivă și crește perioadele de calm și de diminuare a stresului. În cazul insectelor eusociale (foarte sociale), cum ar fi furnicile, albinele, termitele și viespile, care acționează în mod cooperant cu colegii de stup, mecanismul de recunoaștere s-a specializat și oferă o modalitate prin care cei ale căror sarcini includ protecția cuibului să identifice cu ușurință membrii stupului, spre deosebire de invadatorii externi. Aceste forme de recunoaștere intraspecifică sunt atât de importante din perspectiva selecției darwiniene, încât o serie de mecanisme diferite au evoluat pentru a le pune în aplicare folosind o varietate de procese perceptive. Unele implică semnale olfactive (multe specii acvatice), altele recrutează stimuli auditivi (păsări cântătoare), altele configurație corporală (insecte) și afișaj vizual (multe mamifere și specii aviare).

Ceea ce vedem aici este un mecanism care a evoluat la aproape toate speciile, iar atunci când biologii evolutivi observă acest tip de universalitate, devin foarte interesați. Atunci când o strategie de selecție apare la o gamă largă de specii într-o varietate de circumstanțe, puteți fi siguri că nu este vorba despre o tactică banală, de tipul „lovește sau nu lovește”. Aceste tipuri de reacții naturale sunt adânc înrădăcinate în genomul fiecărei specii. Ele sunt foarte vechi din punct de vedere evolutiv și funcționează în mare parte în afara conștiinței. Nu „credem” că i-am reperat pe ceilalți oameni într-o pădure, printre o trupă de cimpanzei, așa cum cimpanzeii nu „cred” că au făcut distincția între noi și partenerii lor intraspecifici.[24] Este un rezultat al sistemului nostru foarte eficient de detectare în interiorul speciei, iar ceea ce ne oferă este o prezumție automată, de rutină, că orice conspecific identificat va prezenta caracteristici și va avea comportamente care sunt în conformitate cu ale noastre. În consecință, atunci când ne confruntăm cu o inteligență artificială sau cu un robot care are aceste caracteristici familiare, această atracție se declanșează și ne trezim că acordăm artefactului pe bază de siliciu o conștiință asemănătoare celei umane, o minte. Această tendință primește un alt impuls emoțional de la sistemul la fel de larg răspândit care guvernează capacitatea de a constata că „ceva/cineva este acolo” - detectarea agenților.

Detectarea agenților.

O componentă a capacității bine puse la punct de a detecta conspecificii este capacitatea de a detecta agenții. La urma urmei, fiecare caz de conspecific (sau nu) implică determinarea faptului că a existat ceva/cineva acolo. În acest sens, termenul „agent” se referă la un organism individual, la o entitate, și are conotația că acest agent este capabil de acte și comportamente semnificative - pe scurt, că are acțiune. Este extrem de adaptiv pentru un organism să aibă un mecanism de detectare a agentului bine pus la punct, dacă nu chiar hipersensibil. Dacă sunteți o specie pradă, orice semnal, un zgomot, un foșnet în frunze, o schimbare în tiparele de lumină este cel mai bine interpretat ca un agent potențial periculos. Dacă nu luați măsurile adecvate, ați putea foarte bine să deveniți gustarea

24 Acest aspect inconștient sau implicit al mecanismelor străvechi din punct de vedere evolutiv este un element important într-o teorie generală a cogniției dezvoltată cu câțiva ani în urmă (a se vedea Reber, 1967.1989.1993) și va juca un rol în aprecierea domeniului de aplicare al modelului CBC al conștiinței. Legăturile dintre acestea vor fi dezvoltate mai târziu, în capitolul 5.

de culcare a cuiva şi o fundătură darwiniană. Dacă sunteţi o specie prădătoare, ar fi bine să vă dezvoltaţi un detector la fel de bine acordat. Dacă nu faci acest lucru, este posibil să îţi petreci zilele flămând şi să pierzi în faţa altora ale căror mecanisme de detecţie mai ascuţite le-au permis să prindă toate prăzile gustoase, să îţi mănânce literalmente prânzul.[25]

După cum vom sublinia în capitolele următoare, una dintre caracteristicile cheie ale evoluţiei este că, de obicei, aceasta nu renunţă la formele şi funcţiile utile. Odată ce un mecanism este în funcţiune şi codificat într-o formă stabilă în ADN-ul speciei, acesta tinde să rămână. Pe măsură ce alte specii evoluează şi se ramifică, trăsătura subiacentă le însoţeşte. Aspectul genomului care codifică acest caracter sau această formă poate suferi mutaţii şi poate căpăta alte caracteristici; procese mai elaborate pot evolua folosindu-l ca bază, dar nucleul rămâne ca un fel de platformă pe care se dezvoltă trăsăturile apărute ulterior. Deoarece sunt mecanisme extrem de adaptative, atât detectarea agenţiilor, cât şi recunoaşterea conspecifică au fost identificate la începutul evoluţiei, sunt în esenţă universale şi se manifestă într-o serie de forme, în funcţie de biosfera în care a evoluat specia.

În consecinţă, atracţia noastră afectivă este evocată în modul cel mai emoţionant de agenţii care imită forma umană şi simulează mişcarea şi reactivitatea umană întruchipată. Un robot care face ceva este văzut imediat ca o entitate distinctă, un agent; dacă seamănă cu noi, va declanşa procesul de identificare conspecifică. Cu cât un robot arată şi se mişcă mai mult ca un om, cu atât mai uşor îi vom acorda trăsături umane, gândire umană şi conştiinţă umană. Kate Darling, robotician la Massachusetts Institute of Technology (MIT), care a venit în acest domeniu după o carieră anterioară în drept şi etică, a examinat condiţiile în care oamenii dau roboţilor stări interne şi manifestă empatie faţă de aceştia. Acestea includ,

> *- <u>atribuirea de nume</u> - aşa cum am menţionat mai devreme, chiar şi oamenii de ştiinţă, cercetătorii din domeniul inteligenţei artificiale, îşi numesc de obicei roboţii,*
> *- <u>fizicalitatea</u> - robotul există în spaţiul real, nu pe un ecran sau într-o înregistrare video,*
> *- <u>mişcare autonomă percepută</u> - observatorii simt că AI-ul ia decizii cu privire la acţiuni şi,*
> *- <u>comportamentul social</u> - robotul are trăsături asemănătoare celor umane, cum ar fi ochii şi faţa, şi este programat să se angajeze în ceea ce pare a fi moduri de interacţiune umane*.

Darling şi alţii din domeniu iau în serios aceste probleme şi analizează dacă protocoalele standard care guvernează comportamentul etic faţă de speciile sensibile ar trebui extinse la roboţii „sociali” care evocă aceste răspunsuri emoţionale. Mai multe ţări au deja comisii care analizează această problemă, inclusiv Coreea de Sud, Japonia şi Uniunea Europeană. Argumentul nu este că trebuie să protejăm roboţii de suferinţă mai mult decât ar trebui să protejăm uscătoarele de păr sau toaletele cu apă[26], ci că tendinţa de a le conferi sentimente şi sensibilitate este atât de puternică încât, odată ce a fost evocată, oamenii reacţionează cu alarmă şi consternare dacă robotul „suferă” într-un fel sau altul[27]. De fapt, după cum a raportat Darling într-un studiu realizat cu câţiva ani în

---

25 Detectarea agenţiei a fost, de asemenea, implicată în apariţia religiei. Antropologi precum Pascal Boyer susţin că credinţele şi comportamentele religioase pot fi privite ca produse secundare ale acestui mecanism şi ale altor mecanisme vechi din punct de vedere evolutiv. Ara Noranzayan, Joseph Henrich şi colegii lor au elaborat teorii privind modul în care societăţile la scară largă au apărut din grupuri tribale mici şi modul în care religiile prosociale s-au dezvoltat pe baza unor procese care îşi au originea evolutivă în funcţii precum capacitatea de a detecta agenţi, de a identifica conspecifici şi de a distinge rudele de cei care nu sunt rude. Acestea sunt lucruri fascinante şi vă recomand cu tărie să citiţi Boyer (2003) şi Noranzayan et al. (2016) pentru o viziune nouă şi provocatoare asupra originilor religiei.

26 Deşi psihologii cognitivi Nicolas Spatola şi Karolina Urbanska au sugerat recent că oamenii ar trebui să ia în considerare posibilitatea de a acorda statut juridic roboţilor în anticiparea apariţiei conştiinţei.

27 Filmul Westworld din 1973 (şi reluarea sa televizată) prezintă roboţi cu aspect complet uman. Cu toate acestea, problemele etice sunt explorate dintr-o perspectivă destul de diferită. Premisa este că oamenii sunt liberi să se comporte în orice mod doresc faţă de ei. Ei pot fi buni, chiar eroici, dar sunt, de asemenea, liberi să violeze, să tortureze sau să ucidă fără repercusiuni. După cum au remarcat psihologul Paul Bloom şi neurologul Sam Harris într-un eseu recent, această platformă teatrală şi reacţiile noastre la ea reprezintă o invitaţie la explorarea acestor şi a unei serii de alte probleme etice esenţiale pentru ceea ce înseamnă să fii om.

urmă, oamenii își manifestă suferința chiar și prin simplul fapt că închid robotul. Pe scurt, tendința de a considera inteligența artificială ca fiind potențial sensibilă se datorează unei serii de factori, de la cei care se bazează pe mecanisme evolutive străvechi la procese sociale bazate pe proiecții antropomorfe.

Rezumat

Permiteți-mi să închei acest capitol cu două puncte-cheie. Primul este afirmația că sensibilitatea, conștiința, conștientizarea - oricum am dori să le numim - nu vor fi instanțiate într-un sistem artificial. În acest proces, am început să dezvolt un argument paralel conform căruia aceste experiențe mentale bazate pe elemente organice nu sunt experiențele privilegiate ale oamenilor, ci sunt, de fapt, întâlnite la cele mai simple specii.

Anticipez o anumită rezistență la această caracteristică a modelului **CBC**. Mă aștept ca mulți cititori să fie de acord cu argumentele privind lipsa de conștiință a roboților, dar s-ar putea să se simtă puțin stânjeniți de teza de bază, deoarece aceasta se bazează pe argumentul că nu există nimic special în legătură cu conștiința umană, nici partea „umană", nici elementul „conștiință". Cei mai mulți oameni, oameni de știință, filozofi și profani deopotrivă, găsesc adesea acest tip de noțiuni mecaniciste tulburătoare. După cum au constatat recent psihologii de la Berkeley, Sara Gottlieb și Tania Lombrozo, în cadrul unei serii de experimente, simpla afirmație că știința poate înțelege cu adevărat fenomene psihologice precum empatia, introspecția, aprecierea estetică, conștiința și liberul arbitru îi face pe oameni să se simtă foarte incomod.

În al doilea rând, propunând principiile de bază ale modelului **CBC**, m-am alăturat altor „simpliști", celor care au susținut că mințile noastre, conștiința noastră autoreflexivă, sentimentul nostru pasional de a fi speciali, poate chiar sufletele noastre sunt doar rezultatul unor procese biologice naturale (Darwin), sau al unor pasiuni și impulsuri primitive incontrolabile (Freud), sau al unor comportamente învățate și condiționate (Pavlov, Skinner), sau al unor calcule mentale reci, fără afecte (Putnam, Hofstadter). **CBC** este, într-adevăr, o teorie pur fizicalistă. Toate mecanismele explicative vor fi legate de principiile fundamentale ale biologiei evolutive. Toate experiențele fenomenologice vor fi legate de operațiile biofizice, neurochimice, toată sensibilitatea va fi urmărită până la procesele de bază intra - și intercelulare. Dacă acest lucru este deranjant, rece, lipsit de sentimente, ei bine, este în regulă, cred... Uneori trebuie să rupi un trandafir în bucăți, petală cu petală, pentru a descoperi baza frumuseții sale. Eu văd acest efort ca fiind vibrant, bogat în potențial și profund incitant, o oportunitate de a privi altfel natura plină de viață a existenței noastre. Dar înțeleg de ce unii nu ar face-o.

Există o veche glumă filozofică care merită povestită aici înainte de a trece mai departe. Un tânăr în căutarea iluminării merge la coliba unui șaman înțelept și bătrân pentru a-l întreba cum poate fi lumea sigură, stabilă. „Ce o ține în loc?", întreabă el. „De ce nu zboară sălbatic în cosmos?" Bătrânul zâmbește și spune: „Pentru că, fiule, stă în siguranță pe spatele unei țestoase mari". Tânărul pleacă dând din cap, ajunge la câțiva metri în afara colibei șamanului și se oprește. Se întoarce și spune: „Scuzați-mă, domnule învățat, dar acest lucru nu este satisfăcător. Pe ce se odihnește marea țestoasă?" Șamanul zâmbește și spune: „Pe spatele unei țestoase și mai mari". Tânărul pare confuz. „Dar, sire, cu siguranță acea țestoasă are încă o problemă, nu-i așa?" „Uite ce e, tinere șmecher", mârâie șamanul, "Sunt doar țestoase până la capăt!" Și așa vor fi mecanismele microbiologice de bază... până la capăt.

2. De ce nu a fost propus până acum un model de tip bază celulară a conștiinței?

Simpla idee că ceea ce ne place să numim „minte” este o caracteristică inerentă a vieții, că sensibilitatea există cu deplină integritate ontologică în cele mai simple organisme, este considerată de mulți, atât în cadrul comunității științifice, cât și în rândul publicului larg, drept o propunere radicală. Și, într-adevăr, așa am simțit și eu odată. Mi se părea absurd ca o creatură unicelulară să fie tratată ca o ființă sensibilă. Acum mi se pare un adevăr atât de evident, încât nu-mi pot imagina să mă gândesc la subiectivitate sau minte în alt mod. A fost nevoie de o întâlnire cu o omidă atrăgătoare în urmă cu mulți ani pentru a mă împinge peste obstacolul conceptual.

Omida mea propedeutică

În primăvara anului 1993, priveam o mică omidă verde (a cărei verișoară ați întâlnit-o deja în figura P.1) ronțăind una dintre plantele mele de busuioc - una dintre puținele lucruri pe care le plantasem în acel an și care nu cedase în fața bolilor sau a neglijenței. Fiind transformat într-unul dintre acele suflete ușoare care nu se simt îndreptățite să stingă viața, am rezistat impulsului inițial de a strivi micul vierme flămând. În schimb, am trecut în modul observator naturalist. Cu cât îmi priveam mai mult „viermele de măsurat”, cu atât mișcările sale deveneau mai fascinante. Desigur, acestea erau rutinizate, repetitive și limitate. Nu existau rugăciuni către zeul busuiocului, nu se alăturau conspecificilor pentru a explora noi modalități de recoltare a recoltei, nu construiau unelte pentru a grăbi munca, nu aveau un sentiment evident de sensibilitate, reflecție mentală, conștiință de sine.

Cu toate acestea, nici comportamentul său nu era chiar atât de mecanic - în sensul de a fi bine structurat, automat, invariabil. Își ridica capul după fiecare mușcătură, îl mișca în jurul său, parcă cernând curenții de aer, verificând crestele de pe marginea frunzei înainte de a se apleca grațios și de a ronțăi un alt semicerc prețios de iarbă. Ar putea fi văzut cu ușurință ca angajându-se într-un proces asemănător unei alegeri - determinând unde să mestece în continuare, cât de mare este bucata pe care o va lua, evaluând dacă este sigur să mănânce acum sau dacă ar fi mai potrivită o altă verificare pentru prădători. M-am uitat în jur la celelalte frunze mușcate mai devreme. Fiecare era tăiată cam în același mod, așa cum ne-am putea aștepta de la un proces mecanic, dar fiecare era destul de diferită, ca și cum ar fi fost luate decizii noi pe măsură ce omida își croia drum de la o frunză la alta - sau cel puțin, factori de mediu noi au determinat comportamente diferite de la o frunză la alta.

M-am trezit că am gânduri eretice și mi-am amintit de avertismentele severe lansate cu peste un secol în urmă de psihologul britanic Conwy Lloyd Morgan cu privire la nevoia psihologiei de a examina adecvarea explicativă a răspunsurilor simple înainte de a le lua în considerare pe cele complexe. Nu am vrut să înzestrez această creatură, a cărei rețea neuronală primitivă abia poate fi calificată drept creier, cu o capacitate operațională mai mare decât cea pe care o deținea în realitate. Dar, totuși, era ceva în acțiunile sale care, pentru a folosi o expresie favorită a cognitivistului Edward Chace Tolman, „mirosea a scop”. Ceva din observațiile mele din timpul celor aproximativ 10 minute mi-a dat sentimentul că repertoriul comportamental al acestei creaturi, deși cu siguranță limitat, nu era produsul unui automat. Era, desigur, o mașină, la fel ca noi, o mașină simplă din punct de vedere fiziologic, neurologic și cognitiv, dar, la fel ca noi, una destul de interesantă.

În timp ce continuam să privesc cum îmi consumă busuiocul, chiar și cu frânele rațiunii aplicate, chestia aia blestemată părea să aibă un fel de conștiință primitivă a lumii din jurul ei. Părea capabil să deosebească sinele de non-eu, el de celălalt, eu de tine. Am întins degetul, iar ea a pășit ușor, ridicându-și capul și părând să mă privească în ochi. Mi-am mutat degetul înapoi spre o frunză, iar ea, cuminte, s-a întors la prânzul ei. Pe scurt, se comporta ca un dispozitiv biologic care

era produsul unei istorii evolutive extinse, aceeași care avea să ne producă pe mine și pe tine. Comportamentul său a arătat tipuri familiare de oscilații și indecizii, tipuri organice de mișcări primare și secundare în care a doua deseori a anulat-o pe prima - ridicându-și capul puțin în sus, apoi mai sus, apoi din nou puțin în jos, apoi în sus și în cele din urmă în jos pentru a rupe o altă bucată de frunză.

În acel moment, părea absolut evident, dincolo de orice îndoială, că omida mea poseda o viață mentală ontologic valabilă, o instanțiere clară a unei conștiințe, o sensibilitate existențială, fenomenologică, internă, esențialmente subiectivă - chiar dacă era foarte simplă și nu făcea mare lucru în afară de a ghida distrugerea culturii mele de busuioc. În clipa următoare, am fost la fel de sigur că această subiectivitate nu ar fi putut apărea în mod magic în acest *proto-Lepidoptera* sau în orice alt animal pluricelular mai puțin complex din punct de vedere biologic. Trebuia să fi existat dintotdeauna, încă de la începuturile vieții.

Și astfel a apărut primul principiu al modelului *Bazei Celulare a Conștiinței* (**CBC**). Dar, dacă această ipoteză era atât de evidentă, de ce nu a fost propusă înainte? De ce nu este un principiu fundamental al biologiei evolutive? Cu siguranță că nu am dat peste vreo înțelegere profundă pe care alții au ratat-o - în special pentru că, după cum vom vedea, abordarea **CBC** ne permite, de asemenea, să reconfigurăm mai multe probleme clasice din filosofia minții, inclusiv, în mod important, o rezolvare a problemei dificile. După cum am descoperit, mai multe propuneri din alte surse s-au apropiat și le vom examina, dar puține par să fi făcut acest ultim pas, iar cele care au făcut acest lucru pur și simplu au punctat și au trecut mai departe fără a urmări implicațiile - și cred că există cel puțin trei motive destul de clare pentru aceasta.

Trei motive de ezitare

Motivul I:

Preocupări cu privire la respectarea principiilor științifice de bază. Din punct de vedere istoric, noțiunea de conștiință non-umană a apărut, nu într-un context științific, ci în mit și religie. Doctrinele animiste și diversele mituri ale creației au, de obicei, o calitate asemănătoare minții investită în toate lucrurile (mai multe despre acest principiu mai jos). Aceste tipuri de credințe misteriene pot fi descurajante dacă cineva dorește să dezvolte o teorie bazată pe biologia evolutivă. Bănuiesc că dorința de a nu fi asociat cu aceste perspective teologice a jucat un rol.

Nu am văzut prea multe în literatura de specialitate care să citeze acest lucru ca factor, dar bănuiesc că a fost și poate fi încă. Eu, personal, sunt sensibil la aceste tipuri de probleme de vinovăție prin asociere. La mijlocul anilor 1960, când am publicat pentru prima dată constatările mele privind învățarea care avea loc în afara intențiilor de a învăța și fără a fi conștient de ceea ce se învăța (a se vedea Reber, 1967), am numit procesul *învățare implicită*. Atunci când colegii m-au întrebat de ce nu i-am spus pur și simplu *învățare inconștientă*, având în vedere că asta era, am explicat că nu am vrut ca munca mea să fie asociată în niciun fel cu psihanaliza freudiană, pe atunci cadrul dominant în care mecanismele inconștiente erau considerate ca jucând un rol în psihologie. Am vrut să stau cât mai departe posibil de ceea ce am considerat a fi o tradiție neștiințifică. Vedeți paralela.

Motivul II:

Preocupări legate de demitologizarea minții. Mai important, bănuiesc, este faptul că oamenii de știință, cu toată obiectivitatea și abordarea lor lipsită de pasiune față de subiectele de cercetare, nu pot scăpa complet din ghearele **memelor** și obiceiurilor culturii noastre - iar unul dintre acestea este că mintea umană are ceva cu adevărat special.

Suntem ființe conștiente, gânditoare, cu talente mentale remarcabile. Suntem singura specie cu un limbaj complet dezvoltat. Ne putem auto-reflecta cu ușurință și ne putem angaja în acte metacognitive precum gândirea despre faptul că gândim și scrierea de propoziții despre scrierea de propoziții. Nu numai că avem conștiință, dar știm că o avem și discutăm constant despre formele și funcțiile acesteia. Putem contempla viitorul, inclusiv imaginându-ne propria moarte, și putem avea experiențe transcendentale, medita, crea și compune. Este greu să nu fii impresionat de noi și este și mai greu să ne imaginăm dotarea altor specii mai puțin complexe și mai puțin sofisticate din punct de vedere cognitiv cu astfel de talente și abilități, chiar și în formă incipientă. Acest tip de protecționism al speciilor există de foarte mult timp și este, bănuiesc, un factor important în motivul pentru care ceva precum imaginea prezentată de teoria **CBC** ar putea fi considerat radical. Înzestrarea tuturor speciilor cu sensibilitate demitologizează conștiința umană - unii ar putea spune că o banalizează - și, în consecință, este probabil să fie condamnată de către colegi, criticată de către colegii noștri științifici și batjocorită de către publicul profan cu un nivel intelectual ridicat.

Aceasta nu este o îngrijorare ipotetică. Este exact ceea ce s-a întâmplat atunci când biologul Donald Griffin a înaintat, nu cu mult timp în urmă, o propunere privind tipurile de experiențe mentale pe care le-ar putea avea non-oamenii. Merită să aruncăm o privire rapidă asupra istoriei; aceasta implică o abordare a studiului animalelor, ceea ce a ajuns să fie cunoscut sub diverse denumiri: ecologie cognitivă, etologie cognitivă sau biologie cognitivă.

Ecologia cognitivă este o ramură a științelor biologice care a apărut prin amalgamarea mai multor abordări ale studiului animalelor, inclusiv zoologia, biologia evolutivă și psihologia comparativă. Aceasta s-a bazat pe o abordare anterioară, cunoscută sub numele de etologie, studiul științific al comportamentului animalelor atât în condiții naturale, cât și în medii controlate. Etologia s-a impus ca disciplină distinctă în anii 1930, odată cu activitatea biologilor Karl von Frisch, Nikolaas Tinbergen și Konrad Lorenz.[28] Abordarea lor se înscria în tradiția behavioristă dominantă la acea vreme. Von Frisch a studiat albinele și a fost primul care a recunoscut rolul dansului ondulatoriu pe care albinele care se întorceau să caute hrană îl foloseau pentru a comunica partenerilor de stup locația plantelor înflorite. Tinbergen a promovat studiul instinctului, dezvoltând un model bazat pe rezervoare de energie ipotetice care, în opinia sa, urmau principii similare cu cele ale sistemelor hidraulice dinamice. Lorenz a studiat amprentarea la gâște. El a devenit cunoscut pe scară largă atunci când filmele cu el înotând alături de un cârd de gâște cenușii nou eclozate care își luaseră amprenta asupra sa au devenit, cum am putea spune astăzi, „virale”.

Dar acești biologi, precum și studenții și colegii lor, au rămas cu toții în limitele metodologice și paradigmatice ale behaviorismului. Datele erau toate măsuri obiective bazate pe răspunsuri observabile. Modelele au fost dezvoltate pe baza unor ipoteze riguroase privind sistemele biologice și genetica. Teoriile instinctului, atât ale lui Lorenz, cât și ale lui Tinbergen, foloseau principii din fizică ca analogii și considerau tendințele instinctuale ca un fel de sistem energetic, creierul fiind tratat ca un rezervor. Nu au existat incursiuni sau speculații cu privire la funcțiile cognitive; niciun pui de căprioară nu a decis să își urmeze mama sau pe Lorenz. Comportamentul era gravat în piatră darwiniană.

Griffin s-a revoltat împotriva acestei tradiții, mai întâi la Harvard și mai târziu la Universitatea Rockefeller din New York. Ideea principală a abordării lui Griffin a fost că animalele, altele decât oamenii, aveau o capacitate cognitivă mult mai mare și mai multe abilități bazate pe cogniție decât se recunoștea anterior. Griffin a avut, de asemenea, îndrăzneala de a susține că aceste repertorii comportamentale sofisticate implicau faptul că multe animale erau conștiente, aveau minte și se puteau angaja în gândire deliberativă.[29]

Propunerile lui Griffin nu au fost întâmpinate tocmai cu bannere fluturând sau cu laude efuzive. *The Question of Animal Awareness* a fost publicată în 1976, când hegemonia behavioristă a deceniilor anterioare era în declin, iar cei care se agățau de ortodoxia conform căreia animalele nu

28 Aproximativ 40 de ani mai târziu, contribuțiile lor au fost recunoscute atunci când cei trei oameni de știință au împărțit Premiul Nobel pentru Fiziologie sau Medicină în 1973.

29 A se vedea Griffin, 1976, 1992 pentru detalii și biografia viitoare a psihologului comparativ Carolyn Risteau pentru informații despre cariera lui Griffin și o analiză detaliată a modului în care munca sa a fost primită de colegii săi.

erau decât niște mașini de stimulare-răspuns simțeau o schimbare de tendință - și deveneau un pic nervoși. Biologul Marc Bekoff a luat notă de rezistența față de Griffin. El și colegul său Colin Allen au identificat, într-un mod amuzant, trei grupuri vag definite de cercetători ale căror lucrări au fost afectate de trecerea etologiei la cognitivism și au descris modul în care aceștia au reacționat la aceasta. Aceștia au fost supranumiți „ucigașii", care l-au atacat pe Griffin și pe cei care au urmat pe urmele sale cu un brici bine ascuțit, hotărât să submineze în totalitate eforturile de cercetare, „scepticii", care au criticat dur, deși erau dispuși să recunoască faptul că unele dintre constatări erau legitime și, eventual, aveau valoare științifică, și un grup mai mic de „susținători", care au salutat noua abordare.

Atacurile Vânătorilor au fost de o natură care putea fi mai mult decât intimidantă. Iată câteva exemple:[30] Biologul britanic John Kennedy[31] a intrat în toată mișcarea ecologiei cognitive:

Dacă oamenii de știință, cel puțin în cele din urmă, încetează să plece de la premisa conștientă sau inconștientă că animalele au minte, atunci este de așteptat ca consecințele să depășească limitele studiului comportamentului animalelor. Dacă vechea problemă minte-corp ajunge să fie considerată ca fiind una exclusiv umană, în loc să fie extinsă la infinit prin regnul animal, atunci și această problemă va fi mai aproape de o soluție.

Psihologul experimental Cecilia Heyes a criticat:

Încercările lui Griffin ... de a îmbogăți studiul comportamentului animal prin cunoașterea conștiinței ... merită atenție ca exemplu al gândirii confuze care este tot mai influentă în acest domeniu.

Filozofa darwiniană Helena Cronin a contribuit cu o recenzie a lucrării lui Griffin:

[Imaginați-vă consternarea din rândul etologilor comportamentaliști atunci când dl Griffin a apărut în urmă cu un deceniu, cu „Problema conștiinței animalelor", ca <u>un sentimental îngăduitor</u>. . . Pentru dl Griffin, toate acestea [inteligența animalelor] sugerează conștiința. El se înșeală. Dacă o astfel de inteligență ar fi suficientă pentru a demonstra conștiința, oamenii de știință și-ar putea face treaba la o cafea, iar filozofii și-ar fi putut împacheta aparatura savantă cu ani în urmă. (Accent adăugat.)

Și un adevărat „zinger" din partea lui Garland Allen, ale cărui opinii clar negative cu privire la programul lui Griffin au trecut în politică:

[Griffin] fără proceduri pentru a ajunge la un răspuns riguros sau științific ... face ca întreaga întreprindere să fie parte integrantă a unei dezvoltări sociale fasciste (deși, în acest stadiu, protofasciste). Având în vedere problemele metodologice inerente domeniului conștiinței animale (precum și sociobiologiei), abordarea problemei în modul exemplificat de cercetători precum Griffin nu poate duce decât la tipul de confuzie și înțelegere falsă a naturii biologice a ființelor umane, pe care se poate construi viitorul fascismului.

Perspectiva lui Griffin a fost, în esență, un efort onest de a îmbogăți peisajul științific. Văzând acest tip de atacuri la adresa sa (și a lui), ar putea, cu siguranță, să dea de gândit oricui contemple noțiuni radicale precum modelul **CBC**. Mai devreme am remarcat reacția unui grup de

30 Le mulțumesc lui Bekoff și Allen pentru exegeza lor atentă a războaielor ecologiei cognitive și pentru că au umblat prin scrierile Vânătorilor pentru a descoperi aceste bijuterii. Mă îndoiesc că le-aș fi putut găsi singur.

31 Kennedy era preocupat de mai mult decât ecologiștii cognitivi. El a atacat utilizarea lor reînnoită a antropomorfismului, despre care credea că a fost supus unei morți bine meritate de către behavioriști. Ca și alții (de exemplu, Rumbaugh, 1994: Burghardt, 2016), am probleme cu poziția lui Kennedy aici. Mai jos, vom analiza mai îndeaproape antropomorfismul și rolurile pozitive pe care le poate juca, rămânând în același timp fidel liniilor directoare ale „științei bune".

participanți la conferință la aluzia ușor deghizată a lui Denis Bray că poate, doar poate, speciile unicelulare ar putea avea experiențe subiective care sunt vag asemănătoare minții. Ca și în cazul asocierii cu teologia și misticismul, nu știu cât de mult, dacă a avut vreun impact, ar fi putut avea acest tip de atmosferă neprietenoasă. Teama de ridicol ar fi putut inhiba pe cineva să extindă noțiunile de sensibilitate și subiectivitate la speciile unicelulare? Nimeni nu știe cu certitudine, dar am văzut de mai multe ori referindu-se la aceasta ca la o „idee nebunească” și, ceva mai blând, ca la o „idee nu complet nebunească”.

Motivul III:

Poate că este greșit. Cei care au luat în considerare extinderea stărilor mentale sau a sensibilității la cele mai elementare forme de viață pot să fi ezitat pur și simplu pentru că au crezut că ipoteza era prea departe și că, cel mai probabil, era greșită. Bineînțeles, acesta poate fi cazul, dar pe măsură ce vom dezvolta argumentele în favoarea ipotezei, vom vedea, sper, că este probabil corect. Dar îngrijorarea este reală și o recunosc. O euristică generală pentru viață este: „*Nu pune o întrebare dacă nu ești pregătit pentru toate răspunsurile posibile*”. O bună heuristică pentru a face știință este: „*Nu prezenta o propunere decât dacă ești pregătit să ți se demonstreze că greșești*”.

Este important să reținem că multe dintre criticile la adresa lui Griffin și a altor susținători timpurii ai cogniției animale au fost lansate din tabăra behavioristă. Unul dintre pilonii behaviorismului a fost acela că abordarea cognitivă a fost împovărată de un antropomorfism fals, o acordare neîngrădită de trăsături și funcții asemănătoare celor umane altor specii care nu merită. Dar behaviorismul este puțin cu noi astăzi și a devenit evident că nu putem înțelege comportamentele multor specii fără să recunoaștem că se întâmplă ceva în interior. Ca urmare, antropomorfismul se bucură de o nouă viață și, deși este cu siguranță adevărat că, atunci când este făcut prost, lucrurile pot merge foarte prost, atunci când este făcut bine este o strategie de cercetare eficientă.

În apărarea antropomorfismului

Să începem prin a recunoaște că, în nu puține cazuri, eroarea antropomorfică a determinat oamenii de știință bine intenționați - precum și publicul larg - să concluzioneze în mod eronat că speciile non-umane au repertorii cognitive mai degrabă umane. Poate cel mai faimos a fost cazul lui der Kluge Hans (sau Hans cel isteț). Atunci când se spune povestea acestui proiect amuzant și greșit, antropomorfismul nu este de obicei identificat drept vinovat. Mai degrabă este prezentată ca o poveste de avertizare cu privire la necesitatea de a controla factori precum părtinirea experimentatorului sau a observatorului și de a recunoaște impactul semnalizării inconștiente. Dar, fără o demonstrație serioasă de eroare antropomorfică din partea tuturor participanților, cu excepția unui asistent de laborator inteligent pe nume Oskar Pfungst, întreaga afacere nu s-ar fi materializat niciodată.

Hans era unul dintre caii din municipalitatea Elberfeld, Germania, deținuți și dresați de Wilhelm von Osten, un profesor de matematică cu un interes mai mult decât trecător pentru ocultism și frenologie, cunoscut pentru excentricitățile și obiceiurile sale de cârcotaș. Hans și-a făcut apariția la începutul anilor 1900, când teoria evoluției a lui Darwin a stârnit un entuziasm considerabil, iar oamenii de știință (și publicul larg) erau fascinați și căutau cu entuziasm originile evolutive ale abilităților umane. Von Osten susținea că Hans putea să înțeleagă limbajul, să citească, să rezolve probleme aritmetice de bază și să răspundă la întrebări de natură generală, toate acestea bătându-și copita. Nu putea, desigur, dar era un cal destul de inteligent și învățase să urmărească îndeaproape persoana care punea întrebarea. Începea să bată din copită și, detectând schimbările

subtile de respirație și de postură ale persoanelor prezente, știa când a ajuns la răspuns și se oprea. Era foarte impresionant.

Pfungst a demascat celebra șaradă prin introducerea mai multor proceduri de control bazate pe asigurarea că nimeni nu știe ce întrebare a fost pusă. Pfungst a fost un erou improbabil. Era un asistent neplătit în laboratorul lui Carl Stumpf. Stumpf era un profesor foarte respectat la Universitatea din Berlin, căruia i se ceruse să facă parte dintr-un grup de experți pentru a determina veridicitatea afirmațiilor lui von Osten. Deși era un promotor pasionat al psihologiei experimentale, Stumpf nu era interesat în mod deosebit de realizarea efectivă a experimentelor. El i-a încredințat această sarcină lui Pfungst, care a continuat să se imortalizeze în analele psihologiei. În mod ciudat, a părut să dispară de pe scenă la scurt timp după aceea, nefinalizând niciodată o diplomă avansată. Incidentul este atât de celebru încât expresia „efectul Hans cel isteț” a ajuns să denumească situațiile în care observatorii oferă indicii inconștiente care pot controla răspunsurile celorlalți.

Hans nu a fost singura creatură „inteligentă” care s-a remarcat în acele zile. Existau gâște inteligente, porci deștepți și câini psihici - iar unele dintre afirmații erau, cel puțin din perspectiva de astăzi, rizibile. Una dintre ele se referea la un raport conform căruia un câine și-a dat seama cum să deschidă o poartă apăsând pe pârghia care ridica bara ce ținea poarta închisă. Acest răspuns este foarte asemănător cu apăsarea barei de către șobolani, un comportament care a devenit unul dintre cele mai utilizate în deceniile behaviorismului skinnerian, în care un șobolan alb care se confruntă cu o bară care iese din peretele unei cutii Skinner a devenit un simbol al psihologiei experimentale pe animale la fel de mult ca un șobolan care se mișcă prin labirint. Din nefericire, persoana care a raportat realizările acestui câine a concluzionat că animalul nu doar apăsa o bară, ci ajunsese să înțeleagă principiile fizice care stau la baza funcționării unui punct de sprijin și i s-au acordat talente de nivelul lui Arhimede.

Robert Rosenthal, care a editat și comentat demascarea clasică a lui Hans de către Oskar Pfungst (a se vedea Pfungst/Rosenthal, 1965), a citat episodul Clever Hans ca alertând zoologii, etologii și alți cercetători în domeniul animalelor cu privire la impactul unor factori precum așteptările experimentatorului, prejudecățile de confirmare și semnalizarea inconștientă. Am devenit foarte devreme sensibil la aceste probleme. Cercetătorii din generația mea au fost instruiți cu privire la pericolele acestor tipuri de abuzuri interpretative. Am fost instruiți să privim antropomorfismul ca pe un demon seducător de care să ne scuturăm usturoiul obiectivist. Eram încă foarte mult în cadrul hegemoniei behavioriste și am fost avertizați să nu atribuim speciilor cu capacitate afectivă sau cognitivă limitată procese mentale sau senzitive pe care nu le-ar putea avea. După cum s-a menționat, nu a fost vorba doar de sfaturi proaste și nu a fost vorba doar de aventuri nefericite precum cea a lui Hans cel Isteț. În critica ascuțită a antropomorfismului, Kennedy a susținut, în mod interesant, că multe dintre argumentele antropomorfice nepotrivite și, în cele din urmă, înșelătoare au fost aduse inconștient, că păcătoșii științifici nu știau de păcatele pe care le-au comis.

Dar antropomorfismul are și o parte bună, care derivă din ceea ce am discutat mai devreme, și anume că există o parte bună în această tendință aproape irezistibilă de a proiecta caracteristici și trăsături umane asupra altor organisme atunci când comportamentele lor par familiare sau asupra artefactelor dacă acțiunile lor le imită pe cele ale oamenilor. Șmecheria în evitarea utilizării greșite a acestei euristici este să fim vigilenți și să ne ferim de extinderea excesivă, să folosim ceea ce Gordon Burghardt numește antropomorfism critic. Aici nu există alte reguli decât aplicarea studiată a bunului simț.

Regula mea de bază este destul de simplă. Mecanismele cognitive și substraturile neuronale care stau la baza acestora și care permit procese foarte sofisticate, cum ar fi autoreferința, înțelegerea dependențelor imbricate, raționamentul inferențial, analizele multifactoriale și altele asemenea ar trebui abordate cu precauția standard împotriva exagerării antropomorfice. Dar cu cât învățăm mai multe despre capacitățile mentale ale altor specii, cu atât mai mult apreciem valoarea antropomorfizării temperate. Când văd doi câini jucându-se, rostogolindu-se în iarbă și sărind unul

peste celălalt, nu am nicio îndoială că sunt fericiți și că bucuriile lor exprimate intern se află pe un continuum evolutiv cu al meu. O mică omidă verde, care roade o frunză de busuioc, este conștientă de circumstanțele sale, iar sensibilitatea sa reală, deși limitată, se află la un nod al unui fir biologic care ne conduce la mine și la tine. Respectatul primatolog Frans de Waal a susținut că recunoașterea empatiei între specii nu este doar acceptabilă, ci este un instrument important în arsenalul cercetătorilor.

Dar, mai important decât argumentul, datele care arată abilități cognitive uneori uimitoare la specii surprinzătoare s-au acumulat. Psihologii cognitivi, etologii și biologii care lucrează atât în laboratoare controlate, cât și în calitate de observatori ai comportamentului în mediul natural au extins gama de funcții cognitive prezentate de alte specii, reducând din exclusivitatea mult clamată a multor abilități umane. Iată doar câteva dintre cele mai interesante exemple. Când le citiți, vedeți dacă nu simțiți acea atracție antropomorfă care vă invită să simțiți că ceea ce fac și experimentează aceste animale are trăsături comune cu noi, oamenii.

- Maimuțele Rhesus prezintă dovezi convingătoare de metacogniție. Cu alte cuvinte, ele știu ce știu și își modifică comportamentul în mod adaptiv atunci când își dau seama că au nevoie de mai multe informații. În recentul studiu inteligent realizat de Alexandra Rosati și Laurie Santos, maimuțele *rhesus* au privit în timp ce unul dintre cele două tuburi așezate pe sol într-un aranjament în formă de „V” era momit cu mâncare. Uneori puteau vedea care braț al aranjamentului era momit; în alte ocazii, vederea le era blocată. Atunci când au privit în timp ce mâncarea era plasată într-un tub, s-au îndreptat în mod fiabil direct către acel tub. În ocaziile în care vederea le era blocată, se duceau de obicei la intersecția dintre cele două tuburi, de unde puteau privi din același punct de observație în josul ambelor tuburi pentru a determina care dintre ele conținea mâncarea. Practic, ei nu făceau niciodată această mișcare de verificare a ambelor tuburi atunci când știau unde fusese plasată mâncarea. Pe scurt, ei știau ce știau și ce nu știau și făceau ajustările corespunzătoare - o abilitate metacognitivă clară care nu a fost observată anterior la non-umani.

- Lucrările recente cu câini au scos la iveală abilități cognitive și lingvistice surprinzătoare. Rico, un *border collie*, a demonstrat că poate învăța numele unui obiect nou într-o singură încercare și fără instrucțiuni specifice. Rico s-a angajat într-o tehnică cunoscută sub numele de cartografiere rapidă (denumită astfel de Susan Carey și Elsa Bartlett în 1978), despre care se credea anterior că este o abilitate privilegiată a oamenilor. În cartografierea rapidă, un copil învață semnificația unui cuvânt nou la o singură expunere, observând că noul cuvânt denumește un obiect nou. De exemplu, imaginați-vă un copil mic într-un parc cu un părinte. Vizavi de ei se află un iaz cu mai multe rațe (un cuvânt pe care copilul îl cunoaște) și o lebădă (un cuvânt necunoscut). Părintele arată cu degetul și spune: „Oh, uite, acolo este o lebădă”. În aceste condiții, copilul învață noul cuvânt și referentul său într-o singură mișcare simplă. Cartografierea rapidă este un mecanism important în construirea vocabularului unui copil și a fost, până la lucrările lui Juliane Kaminski cu Rico, observată doar la oameni. Într-un studiu, Kaminski și colegii săi au pus șapte obiecte, ale căror nume Rico le știa (avea un vocabular de peste 200 de cuvinte), împreună cu un obiect nou. Lui Rico i s-a spus apoi să „meargă să ia __” unde a fost introdus numele obiectului nou. Rico s-a întors cu obiectul corect în peste 90% din încercări (detalii sunt în Kaminski, Call, & Fischer, 2004).

- Abilitățile lui Rico pălesc în comparație cu cele ale lui Chaser, un alt *border collie*[32] al cărui vocabular a depășit 1 100 de cuvinte și care a demonstrat, de asemenea, sensibilitate la ordinea cuvintelor sau, de fapt, la structura sintactică. De exemplu, când i s-a spus „du mingea la Frisbee” într-un proces și „du Frisbee la minge” în altul, Chaser a efectuat acțiunile corecte în aproape toate

32 Border Collie sunt în general considerați cei mai inteligenți dintre toți caninii (orice ciobănesc galez poate depune mărturie). Există suspiciuni că o genă, CTNND2, care codifică proteine asociate cu dezvoltarea cognitivă la om, face parte din genomul border collie. Până în momentul scrierii acestui articol, juriul nu s-a pronunțat încă cu privire la detaliile acestui factor de expresie genetică.

procesele (a se vedea Pillev & Hinzmann. 2011 pentru detalii privind metodologia și abilitățile lui Chaser).

- Recent, a fost difuzat un videoclip care, deși este neoficial și, prin urmare, nu poate fi reprodus, arată o comunicare intenționată remarcabilă între un rechin balenă, cea mai mare specie de rechin și, de fapt, cel mai mare pește care trăiește astăzi, și un grup de scafandri. În mod normal, rechinii balenă înoată pur și simplu pe lângă scafandri, fără niciun contact. Rechinul se hrănește cu plancton și este inofensiv. Dar de data aceasta rechinul s-a oprit, a înotat până la unul dintre scafandri și i-a prins ușor mâna cu fălcile, trăgându-l mai aproape. Apoi scafandrul și-a dat seama că rechinul încerca să îl apropie pentru a putea vedea că în jurul corpului său era înfășurată o plasă de pescuit mare, care îi tăia carnea și îi împiedica mișcarea aripioarelor. Rechinul a dat apoi drumul la mâna scafandrului și a rămas nemișcat în timp ce trei dintre scafandri tăiau și apoi îndepărtau materialul. Imediat ce a fost scos, rechinul s-a îndepărtat înotând încet. Corpul și aripioarele sale erau adânc tăiate de plasă, care probabil se afla acolo de mult timp.

Acest videoclip a fost văzut de milioane de oameni și aproape toți comentatorii îi laudă pe scafandri și remarcă prietenia rechinului. Dar acest episod arată o activitate cognitivă foarte sofisticată din partea peștelui. În primul rând, a înțeles clar că oamenii îl pot salva. De asemenea, a recunoscut că trebuie să „spună” ceva scafandrilor pentru a le atrage atenția asupra problemei, lucru pe care l-a realizat ținând ușor de mână un scafandru. Apoi și-a modificat în mod conștient modul obișnuit de înot (rechinii trebuie să se miște tot timpul pentru ca apa să le curgă prin branhii) și a rămas nemișcat în timp ce scafandrii lucrau la plasă. A înotat doar atunci când a simțit că a fost îndepărtată. Remarcabil în atât de multe feluri.

- Caracatițele și alte cefalopode din captivitate știu că se află într-un mediu nenatural și adesea iau măsuri împotriva acestuia. Sunt mari artiști ai evadării și se deplasează în mod regulat în alte bazine cu apă în căutarea hranei pe care să o fure. Se străduiesc să contracareze aspecte ale mediului care nu le plac, de exemplu, stropind cu apă oamenii din laborator și, în mai multe cazuri povestite, aruncând un jet de apă asupra unor becuri pe care le considerau inutil de luminoase pentru a le sparge sau scurtcircuita. Biologul Jean Boal a raportat cea mai uimitoare manifestare de enervare. Aceasta își hrănea caracatițele captive cu calamari dezghețați, o hrană hrănitoare, dar mai puțin preferată. Se pare că o caracatiță s-a săturat de această dietă neapetisantă. În loc să mănânce calmarul pe care Boal îl lăsase în acvariu, aceasta a păstrat mâncarea până când Boal s-a apropiat din nou de acvariu, apoi, în timp ce ea privea, s-a deplasat deliberat spre țeava de scurgere și l-a aruncat în acvariu (după cum se menționează în Godfrey-Smith, 2016).

- *Amphioctopus marginatus*, o specie de caracatiță din Pacificul de Sud, a fost observată în mod repetat folosind bucăți rupte din coji de nucă de cocos pentru a-și construi adăposturi. Ele caută și selectează cioburile potrivite, le adună, le transportă pe distanțe considerabile și apoi le asamblează cu grijă pentru a crea un adăpost sigur. Acest comportament este atât de frecvent, încât specia este adesea denumită caracatița „nucă de cocos”.

- Carel ten Cate și colegii săi de la Universitatea Leiden din Țările de Jos au descoperit că speciile aviare învață modele implicit[33] (pentru detalii, a se vedea Chen, van Rossum. & ten Cate, 2015 și ten Cate & Okanoya, 2012). Datele lor din studiile cu cinteze zebră au fost remarcabil de similare cu cele din experimentele privind învățarea implicită a modelelor cu oameni adulți (Reber, 1993) și sugari (a se vedea Gomez & Gerken. 2000; Saffran. Aslin, & Newport, 1996).

33 Adică, așa cum am menționat pe scurt într-o discuție anterioară, fără intenția de a învăța și fără cunoștințe atât despre procesul de învățare, cât și despre ceea ce s-a învățat. Vom analiza din nou învățarea implicită mai târziu. Cadrul biologic evolutiv în care este privită împărtășește caracteristici conceptuale importante cu cel care stă la baza abordării CBC.

- Se ştie că, în sălbăticie, ciorile confecţionează unelte pe care le folosesc pentru a extrage hrana din crăpături sau găuri. Alex Weir şi colaboratorii săi au efectuat un studiu controlat în care Betty, o cioară domesticită din Noua Caledonie care nu observase niciodată utilizarea uneltelor la alte ciori, a demonstrat că îşi confecţionează spontan unelte noi. Lui Betty i s-a arătat mâncare pusă într-o găleată mică şi a urmărit cum găleata a fost lăsată într-un tub. Singurul obiect din jur era un fir de metal. Pasărea a îndoit spontan capătul sârmei într-un cârlig, l-a coborât în tub, a băgat cârligul sub mânerul găleţii şi l-a tras în sus (pentru detalii şi fotografii, a se vedea Weir. Chappell. & Kacelnik, 2002). Acest act remarcabil o face cu siguranţă să pară că Betty, şi probabil toate celelalte ciori din Noua Caledonie, au intenţionalitate, în sensul că acţiunile lor nu au sens decât dacă interpretăm acţiunile ciorii ca fiind îndreptate către anumite obiecte.[34]

- Apoi sunt corbii ale căror abilităţi cognitive şi afective se dovedesc a fi cu adevărat remarcabile. Unele dintre cele mai interesante lucrări recente sunt realizate de omul de ştiinţă cognitiv Mathias Osvath şi echipa sa de la Universitatea Lund din Suedia, împreună cu colegi de la Universitatea din Viena. Într-un studiu recent, aceştia au arătat că corbii pot planifica viitorul. Aceştia au primit jetoane care puteau fi schimbate pentru mâncare şi unelte care puteau fi folosite pentru a deschide o cutie cu momeală. După o aşteptare de 15 minute, corbii au selectat jetonul şi/sau instrumentul potrivit în 80% din încercări şi, în mod remarcabil, când li s-au prezentat aceleaşi opţiuni 17 ore mai târziu, au ales obiectele potrivite în 90% din cazuri. În mod curios, a trebuit să se renunţe la unul dintre corbii din cadrul studiului atunci când acesta a găsit o modalitate de a deschide singur recipientul cu momeală. În cadrul unui test de autocontrol, păsările au demonstrat niveluri observate anterior doar la oameni, renunţând la o recompensă mică atunci când ştiau că o recompensă mai mare va fi disponibilă câteva ore mai târziu (detalii despre aceste experimente sunt disponibile în Kabadavi & Osvath. 2017).

În cadrul unui alt studiu, corbii au demonstrat că nu numai că pot recunoaşte feţe umane individuale, dar pot şi să ţină ranchiună. A fost creat un joc de troc reciproc în care corbii puteau schimba un aliment mai puţin dorit (pâine) cu un aliment preferat (brânză). Schimbul a avut loc cu doi experimentatori diferiţi, dintre care unul a jucat întotdeauna corect şi a predat brânza în schimbul pâinii. Celălalt a jucat de cele mai multe ori corect, dar, într-un singur caz, a înşelat corbul, refuzând nu numai să predea brânza, ci şi mâncând-o. Până la o lună mai târziu, corbii au refuzat să mai joace jocul cu experimentatorul care îi păcălise. A existat chiar un indiciu că unii corbi care au urmărit pur şi simplu această mică dramă nu au avut chiar încredere în persoana care a înşelat un conspecific (pentru detalii, a se vedea Muller et al., 2017).

Uneori, corbii păreau mai inteligenţi decât experimentatorii. În cadrul unui alt studiu, corbii au urmărit cum mâncarea era plasată într-un tub destul de lung, astfel încât corbul să nu poată ajunge la ea. Experimentatorii au pus la dispoziţie un băţ, pe care pasărea îl putea împinge în tub, astfel încât mâncarea să cadă pe partea cealaltă. Un corb a găsit băţul greu de folosit şi l-a aruncat, s-a dus la o grămadă de aşchii de lemn, le-a adus înapoi şi a început să le împingă pe rând în tub până când mâncarea a căzut pe celălalt capăt.

- S-a demonstrat că rățuștele îşi imprimă amprenta nu numai asupra obiectelor fizice, ci şi, în condiţiile potrivite, asupra unor manifestări relaţionale abstracte. Antone Martinho şi Alex Kacelnik, zoologi la Universitatea Oxford, au reuşit să amprenteze rățuștele pe mai multe obiecte stimulente diferite atunci când toate reflectau o caracteristică „acelaşi versus diferit”. Cu alte cuvinte, obiectele

34 Intenționalitatea este o caracteristică a unei stări mentale care are un anumit grad de „despre” ea. Se spune că gândurile care vizează sau reprezintă obiecte, evenimente sau stări de fapt au intenționalitate. Conceptul a stat la baza teoriei timpurii şi influente a minţii a lui Franz Brentano. În opinia sa, intenționalitatea distingea stările mentale de cele fizice. Brentano a fost unul dintre cei mai influenţi gânditori şi profesori de la sfârşitul secolului al XIX-lea. Printre numeroşii săi studenţi s-au numărat Edmund Husserl şi Sigmund Freud. Om de conştiinţă, Brentano s-a pregătit iniţial pentru preoţie şi a fost hirotonit în 1864. Nouă ani mai târziu, s-a retras în semn de protest faţă de dogma infailibilităţii papale. Ani mai târziu, deşi renunţase la preoţie, a renunţat la postul de profesor la Universitatea din Viena pentru a se căsători. În anii 1910 a fost un pacifist declarat.

de stimulare s-au schimbat, dar comportamentul de imprimare a persistat atât timp cât proprietatea abstractă pe care au imprimat-o a fost menținută. Rățuștele imprimate pe „același" au urmărit oricare două obiecte atât timp cât acestea erau identice. Cele cărora li s-a imprimat „diferit" au urmărit doar două obiecte neidentice.

- În aceeași ordine de idei, John Magnotti și colegii săi de la Baylor College of Medicine au descoperit că un spărgător de nuci al lui Clark, un corvid cunoscut pentru capacitățile sale cognitive uneori uimitoare, poate, de asemenea, să învețe un concept abstract același/diferit, în cuvintele lor, „mai bine decât orice altă specie non-umană (inclusiv primatele non-umane) testată până acum pentru aceeași sarcină". Aceste mai multe studii cu corvide și păsări de apă arată dovezi clare că păsările pot reține concepte abstracte. Nu cu mulți ani în urmă, majoritatea psihologilor considerau că astfel de reprezentări nonfizice și abstracte la speciile aviare erau, dacă nu imposibile, extrem de improbabile.

- Biologul Lars Chittka de la Universitatea din Londra și colegii săi au raportat recent că bondarii pot imita acțiuni observate, cum ar fi mutarea unei bile mici în centrul unui cerc mare. Dar fac acest lucru numai atunci când observă acțiunile desfășurate de agenți (albine model simulate), dar nu și dacă bilele se mișcă „magic" (prin magneți ascunși). Albinele lui Chittka au demonstrat, de asemenea, abilități de rezolvare a problemelor, optând pentru cea mai puțin solicitantă dintre alternative. Atunci când li s-au dat mai multe bile din care să aleagă, ele au selectat-o în mod fiabil pe cea mai apropiată de centru și au generalizat acțiunile la bile care aveau o culoare diferită de cele utilizate în formare (detalii în Loukola, Perry și Chittka. 2017). În alte studii, grupul lui Chittka a constatat că bondarii pot număra, pot rezolva un puzzle de tragere a sforii pentru a îndepărta un capac de pe o floare și pot recunoaște fețele umane.

- În 1906, Herbert Jennings, primul biolog celular, a descris comportamentul lui *Stentor roeselii*, un protozoar unicelular, atunci când a fost introdus un iritant. Stentor este în mod normal sesil (staționar) și stă în mod obișnuit atașat de o plantă acvatică sau de detritusurile de pe fundul mării, cu pâlnia sa de admisie în formă de trompetă (în esență, gura sa) acoperită cu cili, deschisă și absorbind particule. Cele care sunt consumabile sunt absorbite. Cele care nu sunt acceptabile sunt respinse printr-o reacție stereotipă în care *Stentor* se apleacă și le scutură. Într-o serie de studii atente, Jennings a introdus mai întâi un nor de particule aversive de colorant roșu în pâlnia lui Stentor. Acesta le-a absorbit, le-a găsit neatractive și a trecut prin ritualul de respingere a lor. Jennings a repetat apoi procesul. Din nou, particulele au fost respinse. Dar, după trei sau patru experiențe, *Stentor* a părut să le recunoască și, în loc să încerce să determine dacă sunt comestibile, și-a contractat corpul trăgând părțile pâlniei în interior. După un minut sau două, s-a desprins și și-a reluat căutarea hranei. Jennings a reintrodus particulele nocive de colorant roșu. *Stentor* a răspuns smulgându-și piciorul de pe pernă și înotând, căutând un loc nou, mai sigur, unde să trăiască. Odată ce l-a găsit, a secretat o picătură de mucus lipicios, s-a reatașat și și-a reluat căutarea hranei.

Comportamentele fascinante ale lui *Stentor* sunt greu de observat fără un ajutor vizual, având în vedere că are doar doi milimetri înălțime. Dar încercați ceva aici, ceva care să stârnească acea atracție antropomorfă. Imaginați-vă acest animal, nu ca pe un mic eucariot unicelular sesil[35], ci ca pe o specie acvatică înaltă de un metru și vedeți dacă nu simțiți atracția antropomorfică, vedeți dacă simțiți dorința de a-l investi cu capacitatea de a simți ceea ce este benefic și de a-l distinge de ceea ce este aversiv, de a învăța din experiență și de a anticipa ce fel de stimuli îi vor ieși în cale, până când, într-un mod care amintește de caracatița lui Boal, se va sătura de acest mediu enervant și își va manifesta nemulțumirea prin mijloace comportamentale evidente. Jennings a simțit la fel.

35 „Eu" în greacă înseamnă «adevărat» sau «real», iar «karyote» se referă la un «nucleu». În biologia celulară, orice specie al cărei material genetic, ADN-ul, este încapsulat într-o membrană nucleară este un eucariot. Noi suntem eucariote.

Spre sfârşitul volumului său din 1906, el a opinat (despre *amoebe*, nu despre *Stentor*, dar ideea este aceeaşi):

Scriitorul este ferm convins, după un studiu îndelungat al comportamentului acestui organism, că dacă Amoeba ar fi un animal mare, astfel încât să intre în experienţa de zi cu zi a fiinţelor umane, comportamentul său ar determina imediat atribuirea unor stări de plăcere şi durere, de foame, dorinţă şi altele asemenea, exact pe aceeaşi bază pe care atribuim aceste lucruri câinelui.

- Bacteriile care trăiesc în ceea ce se numeşte biofilme, constând dintr-un număr mare de indivizi legaţi între ei printr-un fel de lipici organic lipicios, pot comunica între ele atât în interiorul speciei, cât şi între specii, utilizând atât procese biochimice, cât şi electrice. Primul tip de comunicare are loc atunci când bacteriile individuale emit molecule care sincronizează dimensiunea şi activităţile unei colonii prin intermediul unui proces cunoscut sub numele de „*quorum sensing*”, semnalat pentru prima dată în 1970 (a se vedea Nealson, Piatt, & Hastings, 1970). Bacteriile îşi monitorizează mediul înconjurător şi secretă substanţe chimice care semnalează altora informaţii importante despre mediu. De exemplu, dacă celulele determină că nivelul nutrienţilor a scăzut periculos, ele secretă molecule care le spun celorlalte să înceteze diviziunea celulară. Atunci când nivelul nutrienţilor îşi revine, ele transmit un fel de mesaj „clar” şi colonia îşi reia creşterea. Acţiunile bacteriilor individuale ar avea un efect redus sau inexistent, însă acţiunea controlată şi coordonată a mai multor bacterii permite coloniei să facă aceste schimbări adaptative.

Atunci când colectivele de bacterii devin mari, comunicarea trece la una bazată pe eliberarea de ioni de potasiu, care trimit unde de particule încărcate electric prin colectiv. Acest mecanism de eliberare a ionilor pentru transmiterea informaţiilor este utilizat în întregul regn animal. Aceste forme de comunicare între cele mai simple forme de viaţă vor fi discutate mai în detaliu în capitolul 4. În mod evident, acestea vor juca un rol semnificativ în susţinerea empirică a modelului bazei celulare a comunicării.

Un alt element intrigant al acestei forme de comunicare este faptul că ea are loc în circumstanţe în care distincţia dintre detectarea şi reacţia organismelor unicelulare individuale şi cea a speciilor pluricelulare începe să se estompeze. Dacă organismele unicelulare individuale sunt sensibile, trebuie să existe un proces prin care reprezentările lor experienţiale individuale devin asumate de conglomerat, de formele multicelulare care au evoluat. Aceasta este, de asemenea, o problemă importantă pentru **CBC**, pe care o vom analiza mai detaliat în capitolul 5.

Ultimele câteva exemple pe care le-am descris sunt deosebit de interesante deoarece insectele, speciile aviare, peştii şi cefalopodele au sisteme nervoase foarte diferite de cele ale mamiferelor şi, desigur, speciile unicelulare precum *Stentor roeseli* şi bacteriile nu au niciunul. Spre deosebire de mamifere, păsările nu au un neocortex. Există celule care sunt omologii centrelor şi căilor corticale ale mamiferelor şi care au o bază evolutivă comună, dar aceste celule protocorticale se găsesc doar în nuclee din creierul păsărilor. Nu există nicio structură neuronală care să semene cu neocortexul mamiferelor. Capacităţile lor cognitive remarcabile par să provină din funcţiile densităţilor neuronale cerebrale.

După cum au descoperit cercetătoarea în neuroştiinţe Suzana Herculano-Houzel de la Universitatea Vanderbilt şi colegii săi (a se vedea Olkowicz et al., 2016), speciile aviare au neuronii deosebit de bine înghesuiţi în creierul anterior, atât de mult încât densitatea lor le depăşeşte, cu mult, pe cele găsite la mamifere.

Cefalopodele au creiere care sunt şi mai îndepărtate de ale noastre sau, de altfel, de ale tuturor vertebratelor. O caracatiţă are un creier mare, central, prin care trece, în mod ciudat, tractul digestiv, dar arhitectura neuronală generală este mai distribuită decât la vertebrate. De asemenea,

caracatițele au ganglioni împerecheați pe tot corpul și fiecare braț are aproape la fel de mulți neuroni ca și creierul central.

Bondarii au creiere minuscule, cu aproximativ 1 100 000 de neuroni și o arhitectură destul de neobișnuită, pe care tehnicile sofisticate de RMN încep să o dezvăluie.

Bacteriile, desigur, sunt un domeniu preneural,[36] dar sunt capabile să proceseze o cantitate mare de informații și să răspundă în mod adaptiv și adecvat. Vom analiza mai îndeaproape organismele unicelulare mai târziu, deoarece nucleul **CBC** se bazează pe funcții ale acestora care implică sensibilitate.

Pe măsură ce literatura despre cogniția animală crește, fiecare nouă descoperire dezvăluie ceea ce Donald Griffin bănuia în anii 1970, și anume că rădăcinile funcției cognitive umane se găsesc în alte specii, mai simple. Punctul de vedere al lui Griffin, care constituie unul dintre principiile-cheie ale modelului **CBC**, a fost că pur și simplu nu are sens să tratăm funcțiile noastre cognitive și afective bogate ca și cum acestea ar fi apărut brusc, fără să fie solicitate, la oameni. Evoluția nu funcționează așa. Trăsăturile, aptitudinile, funcțiile, procesele și formele evoluează de-a lungul timpului geologic, iar trăsăturile, aptitudinile, funcțiile, procesele și formele lor fundamentale vor fi observate la alte specii din panoplia filogenetică. O poziție antropomorfă bine pusă la punct poate fi un instrument de cercetare important, deoarece permite să se stabilească unde și în ce specii să se caute rădăcinile comportamentelor. Este mult mai complicat să încerci să explici comportamentul altor specii fără să presupui că acestea posedă o versiune evolutiv anterioară a funcției mentale umane vizate.

În cele din urmă, antropomorfismul are un alt avantaj, în mare parte ascuns. Acesta reduce tendința de a trata capacitățile umane ca fiind speciale, distincte într-un fel de cele ale altor specii. Daniel Dennett a identificat vinovatul aici și, în maniera sa creativă patentată, l-a numit „gravitația carteziană” - sau o abordare atrăgătoare, dar greșită, a problemelor minții, care te atrage atunci când (în mod conștient sau nu) combini dualismul lui Descartes cu atracția psihologică irezistibilă a narcisismului *Homo sapient*.

Bănuiesc că această prejudecată este unul dintre motivele pentru care atât de mulți filosofi și neuroștiințifici cognitivi au fost prinși în problema dificilă a lui Chalmers, de ce mulți încă se mai gândesc la inutilitatea încercării de a acoperi „diferența explicativă” dintre materie și minte. După cum vom discuta mai târziu, cei care s-au luptat cu acest tip de probleme au fost prinși în vortexul cartezian și s-au concentrat aproape în întregime asupra experienței fenomenale umane, iar problemele pe care încercau să le rezolve erau definite în întregime în sfera conștiinței umane. Abordarea recomandată de modelul **CBC**, prin acordarea unor forme efective, deși mai circumscrise, de viață mentală și sensibilitate tuturor celorlalte specii, oferă un cadru alternativ și invită la o poziție științifică diferită.

Conștiința umană este specială, dar nu este un tip distinct. Este doar un alt simbol de-a lungul unui continuum de forme de experiență fenomenală. Pentru a înțelege cel mai bine conștiința umană, trebuie să facem o inginerie inversă. Începeți cu formele anterioare care, în contextul modelului **CBC**, vor fi primele specii care au ieșit din nămolul și noroiul primordial. Începeți căutarea trăsăturilor esențiale ale sensibilității, nu la apoteoza sa, ci la începuturile sale pietonale.

O scurtă paranteză despre durerea peștilor - o lecție de obiect

Mai mulți biologi și neuroștiințifici, printre care Brian Key de la Universitatea din Queensland și Stuart Derbyshire de la Universitatea Națională din Singapore, au adoptat poziția extrem de controversată conform căreia speciile de pești, deoarece nu dispun de structurile și căile neuronale care au fost identificate ca fiind esențiale pentru procesarea durerii la oameni, nu simt durerea - în sensul că experiența lor în fața evenimentelor nociceptive reflectă o suferință reală.

36 Clasificarea mai mult sau mai puțin acceptată a vieții identifică trei domenii: Bacteria, Archaea și Eukaryota, acestea din urmă distingându-se prin faptul că au un nucleu încapsulat într-o membrană.

Cred că această propunere este greșită, și nu doar greșită ca o problemă de împărțire lungă, ci fundamental, tragic de greșită. Greșeala sa nu rezultă dintr-un model slab dezvoltat sau dintr-un raționament ilogic sau dintr-o colectare neglijentă a datelor, ci dintr-un eșec de a aprecia tipurile de principii evolutive fundamentale care stau la baza teoriei **CBC**.

Nu sunt singurul care respinge argumentele lui Key și Derbyshire - care sunt împărtășite de alții (a se vedea Rose, 2007). Articolul lui Key a apărut în revista *Animal Sentience* care, așa cum am menționat mai devreme, este o revistă cu evaluare *inter pares* și comentarii deschise și (până în prezent) peste 40 de oameni de știință și filosofi au oferit comentarii. Foarte puțini dintre aceștia au fost de acord cu Key, dar niciunul nu a luat calea pe care teoria **CBC** ar recomanda-o. În critica mea, am arătat cum un pic de antropomorfizare ajută la crearea unui argument împotriva lui Key.

Atunci când întâlnim o creatură rănită, chiar și un pește al cărui strămoș comun se află în trecutul îndepărtat, simțim o conexiune empatică, iar această reacție afectivă este un indiciu că înțelegem valența emotivă comună, nucleul experiențial comun. O stea de mare care își încrețește brațele atunci când este tăiată simte durere și se află pe un continuum cu durerea mea atunci când brațul meu este tăiat. Un pește care se zbate disperat pentru a-și scoate cârligul din gură suferă, iar eu îi recunosc suferința și îi pot identifica legătura comună cu carnea mea. Rechinul balenă pe care l-am descris mai devreme, care suferea din cauza plasei groase înfășurate în jurul corpului său, suferea în mod clar, iar comportamentul său deliberativ față de scafandrii care l-au salvat este revelator.

Poziția Key-Derbyshire-Rose este că, deși aceste comportamente par a reflecta o experiență de durere existențială, ele sunt de fapt doar reacții mecanice la stimularea nociceptorilor și că aceste specii, lipsite de structurile și căile cerebrale necesare pentru a interpreta experiența, nu se află într-o stare subiectivă pe care am numi-o durere. După cum a afirmat Derbyshire în eseul său în apărarea lui Key:

Durerea peștilor nu va implica senzații negative identificabile în mod explicit, generând teamă, îngrijorare și un comportament de protecție organizat în mod conștient. Durerea peștilor va fi mult mai puțin elaborată, implicând ceva perceput, ca o imediatețe, dar nu înțeles, ca o intruziune nedorită în ființa subiectivă a peștilor.

Această poziție este goală din punct de vedere epistemic. Ea redefinește pur și simplu durerea ca fiind ceva ce doar oamenii (sau alte specii cu structurile corticale potrivite) experimentează, deoarece noi, spre deosebire de un pește, putem contempla aspectele aversive ale unei frânghii care ne taie carnea, iar un pește poate doar să o experimenteze. De asemenea, se bazează pe un model de cunoaștere a peștilor care este infirmat de cercetări recente. După cum au demonstrat biologul marin Redouan Bshary și colaboratorii săi, peștii se implică într-o serie de comportamente sociale complexe. Aceștia cooperează (cum ar fi în colaborările de vânătoare între specii[37]), se știe că trișează atunci când pot scăpa, folosesc un limbaj corporal surprinzător de sofisticat pentru comunicare și rezolvă probleme tactice mai bine, în unele cazuri, decât primatele (pentru detalii privind aceste experimente, a se vedea Abbott, 2015). Imaginea peștilor descrisă de grupul lui Bshary este departe de cea rigidă și mecanică pe care o presupune Derbyshire.

Argumentul inițial al lui Key se bazează pe identificarea căilor și centrelor neuronale din creierul uman care moderează și procesează experiența conștientă a durerii fizice. Odată identificate astfel, strategia de cercetare este de a determina ce alte specii au aceste căi și centre neuronale.

37 Grouperii și anghilele moray lucrează adesea împreună. Viermii sunt vânători în ape deschise, iar morenele excelează în spații închise. Bshary a observat numeroase cazuri în care mreana trezește morenele, invitându-le să o urmeze până la o crăpătură în care s-a ascuns un pește pradă. Odată ajunsă acolo, anghila se apucă să ucidă și, în acest fel, aruncă toți ceilalți pești din cavitate, care sunt apoi înghițiți de mreană. Grupul lui Bshary estimează că această vânătoare în cooperare este de cinci ori mai eficientă decât ceea ce poate face fiecare specie pe cont propriu.

Celor care le au, Key și colegii săi sunt dispuși să le acorde o experiență conștientă, dishedonică, asemănătoare cu a noastră. Speciile cărora le lipsesc aceste structuri cerebrale sunt considerate a fi doar entități mecanice insensibile, ale căror reacții sunt determinate de procese genetice fără componente subiective.

În acest cadru, peștele care plutește pe fundul unei bărci cu vâsle prezintă doar o reacție evoluată la o lipsă de oxigen, dar nu simte nimic asemănător cu ceea ce am simți noi dacă ne-am îneca. Dacă are o formă de experiență cu efect negativ, peștele, lipsit de conștiință, nu este conștient de acest lucru - o poziție în contradicție clară cu acțiunile rechinului-balenă menționat mai sus, care a acționat în mod deliberat în moduri despre care știa că vor diminua durerea resimțită. În lumea filosofică a lui Key, dacă nu ești conștient de durerea ta, aceasta nu contează ca durere - din punct de vedere ontologic. Această poziție, din nefericire, nu este neobișnuită. Economistul australian Yew-Kwan Ng, într-un comentariu la lucrarea mea din *Animal Sentience,* a remarcat:

[O] problemă mai importantă este aceea de a ști cu un grad ridicat și crescând de probabilitate care specii de animale sunt capabile de sensibilitate afectivă. Pentru speciile nesimțitoare, nu trebuie să ne pese de bunăstarea lor, deoarece nu au nicio bunăstare.

Deoarece poziția **CBC** este că până și cele mai simple organisme au „sensibilitate afectivă", atunci, în principiu, toate organismele au „bunăstare". Este de datoria noastră să nu provocăm durere sau suferință atunci când acestea pot fi evitate sau nu sunt contrabalansate de alte considerente utilitariste - care sunt multe și despre care vom discuta câteva dintre cele mai semnificative în capitolul 5. Adoptarea perspectivei **CBC** invită la examinarea unui număr de probleme etice care, în mod normal, nu fac parte din discuțiile despre conștiință.

Poziția adoptată de cercetători precum Key și Derbyshire are unele elemente interesante, dar este fatalmente greșită din următoarele motive. Prima dificultate este că ei (în special Key) par să presupună că există un singur mod în care un organism poate experimenta durerea. Funcțiile pot și adesea au apărut de mai multe ori, urmând căi adaptative diferite de fiecare dată. Insectele, păsările și mamiferele au dezvoltat zborul și toate au urmat căi evolutive foarte diferite. Corvidele, cefalopodele și primatele folosesc arhitecturi neuronale foarte diferite pentru a ajunge la utilizarea uneltelor. Faptul că durerea umană are căi și centre neuronale identificabile nu implică în mod logic faptul că aceste structuri sunt singurele care mediază durerea adevărată, fenomenologică, experimentată și neplăcută. Faptul că conștiința umană are caracteristicile sale particulare și fundamentele neurologice asociate nu înseamnă că nu există și alte modalități de instanțiere a unei astfel de stări subiective. După cum au remarcat neuroștiințificii Antonio și Hannah Damasio în comentariul lor privind teza lui Key, există dovezi neuroanatomice și neuropsihologice considerabile care susțin propunerea că stările afective, precum durerea, nu sunt asociate exclusiv cu structuri corticale specifice.

În al doilea rând, teza lui Key presupune că, la un moment dat în istoria evoluției, capacitatea pentru o formă specifică de reprezentare subiectivă a apărut în realitatea ontogenetică, în timp ce, cu un moment cosmic înainte, nicio specie nu o deținea. Poziția lui Key îl forțează să se confrunte cu ceea ce eu numesc *Dilema Emergentistului,* pe care o vom analiza mai în detaliu mai târziu. Din punctul meu de vedere, poziția cea mai plauzibilă este aceea că durerea resimțită de fiecare organism este subiectiv reală și extrem de neplăcută în fiecare caz. Este posibil ca manifestarea sa cea mai complexă să fie la oameni, cu gama noastră adesea creativă și nebunească de moduri de a suferi, dar ea a început cu speciile primitive care experimentau stările subiective într-o formă mai redusă - poate ca ceva de genul unui *quale* negativ neplăcut. În cadrul lui Nagel, există, de fapt, „ceva asemănător" cu a fi un pește care suferă, un păianjen pe o suprafață fierbinte, o anemonă de mare atacată sau o bacterie *Escherichia coli* într-o soluție concentrată de sare.

În al treilea rând, invocând abordarea „critică” a lui Burghardt față de un antropomorfism pragmatic, este clar că, în această dispută particulară, sarcina probei nu revine adversarului, ci susținătorului. Key susține că criticii săi nu au reușit să demonstreze dincolo de orice îndoială rezonabilă că peștii simt, de fapt, durere. Dar aceasta nu este responsabilitatea lor. Comunitatea științifică nu este obligată să accepte că peștii nu simt durerea, deoarece nu au structurile neuronale specifice primatelor. Mai degrabă, pe baza motivelor etice pe care antropomorfismul burghardtian le invită, povara este de partea cealaltă. Peștele care se zvârcolește în cârlig arată ca și cum ar suferi, iar Key are obligația de a prezenta dovezi convingătoare că nu este așa.

După cum a remarcat neurologul Anil Seth, poziția lui Key încalcă și principiul precauției. Mai simplu spus, dacă Key se înșeală, dacă ajungem la concluzia eronată că peștii și alte specii cărora le lipsesc căile neuronale specifice identificate de Key nu suferă atunci când sunt răniți fizic, invităm la un tratament insensibil și crud al marii majorități a speciilor de animale de pe planeta noastră. Comentariul lui Ng citat mai sus este revelator în această privință. Mai târziu, în capitolul 5, vom analiza aceste aspecte, precum și alte probleme legate de etică și bunăstarea animalelor. Angajarea într-o astfel de discuție este o obligație născută din argumentele de bază ale tezei **CBC**.

În cele din urmă, desigur, din punctul de vedere al modelului **CBC**, argumentele lui Key și Derbyshire sunt complet greșite - o concluzie la care ajunge și biologul Victoria Braithwaite în cartea sa din 2010, *Do Fish Feel Pain*. Și este genul de greșeală care rezultă din nerecunoașterea continuității evolutive și a comunității în formă și funcție. Cadrul **CBC** plasează conștiința autentică în fiecare animal viu și asta, evident, include peștii. Key, deloc surprinzător, a avut câteva lucruri extrem de critice de spus despre acea lucrare anterioară a mea (Reber, 2016a), în care am oferit o versiune preliminară a **CBC**.

Cei care sunt, ar putea fi sau poate ar trebui să se simtă confortabil cu celulele simțitoare

Există/au existat destul de mulți filosofi, biologi și specialiști în științe sociale care ar fi putut (ar fi trebuit?) să primească abordarea **CBC** - unii care au prezentat poziții care încorporează propunerea principală într-un cadru mai larg, unii ale căror eforturi de a stabili unde a apărut conștiința în marea amploare a filogenezei și unii care, lucrând în cadrul unei tradiții biologice evolutive, s-au apropiat extrem de mult. Rețineți că această listă nu este nici pe departe exhaustivă. Am ales anumite persoane și poziții dintr-o serie de motive: proeminența teoreticianului, unicitatea și/sau noutatea poziției susținute sau modul în care poziția se raportează la modelul **CBC**.

Charles Darwin.

Probabil ar trebui să începem cu cel mai evident, inițiatorul teoriei evoluției prin selecție naturală. În prima parte a volumului său din 1881, *The Formation of Vegetable Mould through the Action of Worms* (Formarea mucegaiului vegetal prin acțiunea viermilor), Darwin a făcut o serie de observații privind comportamentul viermilor, care l-au condus la concluzia că, oricât de primitivi ar fi, viermii au comportamente surprinzător de complexe și probabil o formă primitivă de sensibilitate. Într-o manieră care amintește de întâlnirea mea cu omida mea, Darwin a abordat mai întâi viermii de pământ ca pe niște creaturi mecanice și lipsite de gândire. El a observat că viermii, deși lipsiți de un sistem vizual, s-ar scufunda totuși în pământ dacă ar fi iluminați brusc. El a presupus, la început, că aceasta era o reacție stereotipă, pur automată, care avea loc fără nicio experiență fenomenală. Abia mai târziu, observații mai atente și mai controlate l-au făcut să își schimbe părerea. El a descoperit că, dacă viermele era distras, atenția sa fiind atrasă de o altă acțiune sau stimul, nu manifesta acest comportament de evadare. Era ca și cum viermele avea o

anumită măsură de control asupra acțiunilor sale sau, cel puțin, poseda o ierarhie de reacții care influența comportamentele pe care era cel mai probabil să le manifeste.

Explorările ulterioare l-au condus la concluzia că viermii pot, de asemenea, să evalueze dimensiunea și masa obiectelor, în special a celor pe care le folosesc pentru a-și astupa intrarea în vizuină. El a mers chiar până acolo încât a ajuns la concluzia că râmele, în cuvintele sale, *„merită să fie numite inteligente*”. Mi-ar plăcea să cred că Darwin, cel mai îndrăzneț gânditor al vremii sale (sau al oricărei alte epoci), ar găsi premisa **CBC** cel puțin demnă de luat în seamă.

Siddhartha Gautama sau, mai simplu, Buddha.

Budismul are atât elemente filosofice, cât și teologice, împreună cu tradițiile, practicile și sistemele sale de credințe. Toate au în comun conceptul de **vijnana**, care este tradus în general prin conștiință, minte sau forță vitală. **Vijnana**, care este discutată în cele mai vechi texte, *suttas* (sau discursuri) pali, este văzută ca o proprietate a tuturor ființelor simțitoare și se bazează pe un set de sisteme specifice, bazate pe simțuri și acțiuni. Buddha o considera forța care dă animație tuturor creaturilor. În acest cadru, pare probabil că Buddha ar fi fost de acord cu abordarea **CBC**.

Animiștii.

Din câte știu eu, nu există animiști în comunitatea științifică din zilele noastre, dar dacă ar exista, aceștia ar întâmpina propunerile **CBC** cu o căscătură plictisită. Animismul este considerat în general drept cea mai veche viziune coerentă asupra lumii. Principiul obligatoriu este că toate lucrurile, vii sau nu, mobile sau nu, toate lucrurile corporale sau nu, toate animalele și plantele, toată respirația, saliva și fecalele, toate otrăvurile, ploile, stâncile, râurile, artefactele, clădirile și chiar lucrurile lipsite de substanță, precum visele și umbrele, sunt impregnate de o calitate animată sau spirituală distinctă. Animismul este larg răspândit. În rândul popoarelor indigene, este atât de răspândit și de integrat în cultura și sistemul de credințe al fiecărui grup încât, așa cum a descoperit antropologul Edward Tylor în secolul al XIX-lea, adesea nu are nici măcar un nume.

Nu este ușor să separăm animismul ca principiu ontologic organizatoric de animismul ca doctrină religioasă, dar, în orice caz, animiștii ar fi cu siguranță de acord cu axioma de bază a **CBC**. Omida mea titulară și organismele unicelulare pe care se concentrează modelul nu numai că sunt binevenite în lumea agenților simțitori de către animiști, dar orice sugestie contrară ar fi considerată absurdă.

Panpsihismul.

Panpsihismul adoptă o poziție destul de radicală, și anume că conștiința sau mintea este un element component al tuturor lucrurilor, vii și nu numai. Ca și animismul, numele școlii de gândire este un indiciu al principiului său de bază. **Pan-** înseamnă toate sau inclusiv toate elementele unui grup, iar **psych**(e) se referă la conștiință sau, ocazional, în versiunile mai teologice, la suflet. Ca și animismul, ideea că cele mai simple organisme sunt înzestrate cu sensibilitate este o consecință naturală a principiului filosofic. Panpsihistii susțin că științele fizice, materiale, oferă doar o perspectivă asupra a ceea ce fac lucrurile de substanță, nu asupra a ceea ce sunt, că, indiferent cât de completă poate fi o descriere a unei acțiuni sau a unui proces, calitățile esențiale subiacente ale elementelor acțiunii sau procesului sunt excluse din analiză.

Panpsihismul este o poziție filozofică veche, dar care, în ultimii ani, s-a bucurat de o renaștere, deoarece mai mulți futuriști, filozofi și neuroștiințifici proeminenți s-au alăturat acesteia. De obicei, este prezentată ca o alternativă la *emergentism*, argumentând că orice poziție care acordă conștiință și minte doar anumitor aspecte ale realității sau anumitor specii trebuie să fie capabilă să prezinte un cadru coerent în care mentalul, fenomenalul, este brusc creat acolo unde nu exista

înainte. Psihologul londonez Max Velmans, care adoptă o viziune puternic panpsihică, a criticat schița timpurie a abordării **CBC** care a apărut în revista *Animal Sentience* (Reber, 2016 a) susținând că nu am mers suficient de departe. Prin neadoptarea unei poziții panpsihice, a susținut el, modelul **CBC** era încă împovărat de *emergentism*.

Desigur, accept această critică, pentru că este valabilă. Dar nu pot să nu remarc că nici Velmans nu este eliberat de această responsabilitate. Și el trebuie să identifice punctul din crearea universului în care a apărut factorul său senzitiv „**pan-**”. Această problemă va fi discutată mai detaliat în capitolele următoare, când vom analiza detaliile modelului **CBC**, dar putem anticipa unde vom ajunge observând că, din punct de vedere logic, nu poate exista nicio poziție care să nu se confrunte cu o versiune sau alta a dilemei *emergentiștilor.*

Cartea lui David Skirbina din 2005, *Panpsychism in the West* (Panpsihismul în Occident), a prezentat o perspectivă istorică, subliniind rolul bogat și influent pe care abordarea generală l-a avut în gândirea occidentală. El a urmărit panpsihismul de la începuturile sale în Grecia antică, unde filosofi precum Heraclit și Thales au abordat versiuni ale acestuia, până la reapariția sa în zilele amețitoare ale secolului al XVII-lea, când a fost îmbrățișat de oameni precum Spinoza și Newton. A continuat cu primele decenii ale psihologiei științifice, când teoreticieni precum Wilhelm Wundt și Gustav Fechner au susținut o variantă care conținea, de asemenea, elemente dualiste clare, cunoscută sub numele de paralelism, în care fiecare eveniment fizic era văzut ca având o experiență mentală particulară, coordonată.[38] În continuare, a examinat filosofia și psihologia contemporană, unde cercetători respectați precum Galen Strawson, David Chalmers, Philip Goff și Max Velmans au îmbrăcat, într-o măsură sau alta, mantia panpsihică pe umerii lor. De atunci, alții, precum neurocercetătorii Giulio Tononi și Christoph Koch, au îmbrățișat-o.[39] Bănuiesc că această reapariție recentă este determinată, în mare parte, de aceleași frustrări care au însoțit lupta asupra problemei dificile. În ciuda eforturilor recente, nu suntem cu mult mai aproape de a răspunde la întrebarea care nu va dispărea („Cum face creierul mintea?”) decât eram pe vremea lui Fechner, iar tensiunea rezultată îi determină pe unii să se îndepărteze de fizicism și materialism și să se îndrepte către abordări mai misterioase.

La prima vedere, este greu să nu te îndepărtezi de panpsihism pur și simplu pentru că pare prea ciudat pentru a fi sensibil sau util ca poziție filosofică. Această manevră retorică a fost anticipată de panpsihicii moderni, precum Philip Goff, și ei sunt sinceri în legătură cu aceasta. Goff a publicat recent un articol scurt, non-tehnic, intitulat „*Panpsihismul este o nebunie, dar este, de asemenea, cel mai probabil adevărat*”.

Problema mea nu este cu „nebunia”. Multe propuneri științifice păreau nebunești atunci când au fost formulate pentru prima dată (deriva continentală a lui Wegener) sau încă mai par nebunești în ciuda tuturor dovezilor empirice de susținere (particulele încâlcite ale lui Einstein, Podolsky și Rosen). Există cel puțin trei probleme profunde cu panpsihismul: Una este legată de modul în care este prezentat ca singura alternativă logică la *emergentism*, ceea ce conduce nu la „nebunie”, ci la incoerență; o a doua este legată de lipsa dovezilor empirice ale sensibilității în locurile în care ar trebui să fie găsită; și, în cele din urmă, principiul „peste tot”. Să le examinăm pe fiecare.

---

38 Fechner a avut, de asemenea, o trăsătură excentrică distinctă în multe dintre scrierile sale, inclusiv publicarea de poezii și fantezii sub pseudonimul Dr. Mises și argumentarea vieții mentale în (unele?) plante. El este, de asemenea, creditat cu fondarea domeniului psihofizicii cu formula sa pentru relația dintre stimulii fizici și experiența psihologică paralelă a acestora ca $p = k \log S/S_0$, unde p este intensitatea percepută a unui stimul S, S_0 reprezintă pragul absolut de detectabilitate, iar k este o constantă de ajustare pentru modalitatea studiată. Cunoscută sub numele de legea lui Fechner sau legea logaritmică, aceasta a avut un parcurs istoric destul de lung, cu sute de experimente care au susținut-o și un număr mai mare care au criticat-o. Fechner și-a încheiat volumul 1860 despre psihofizică cu o glumă amuzantă și profetică potrivit căreia „edificiul său psihofizic” va rămâne în picioare, nu pentru că ar fi neapărat corect, ci pentru că criticii săi nu vor fi capabili să cadă de acord asupra modului de a-l dărâma.

39 Abordarea Tononi-Koch este cea mai dezvoltată abordare a conștiinței care îmbrățișează o formă de panpsihism. O vom analiza mai îndeaproape mai târziu în capitolul 3.

Argumentul emergentismului.

Desigur, singura alternativă aparent logică la una sau alta dintre versiunile panpsihismului (și există multe dintre ele) este una sau alta dintre varietățile *emergentismului*. Dacă conștiința nu pătrunde peste tot ca o consecință naturală a modului în care sunt lucrurile, atunci, în mod obligatoriu, ea trebuie să apară undeva, cândva, în anumite condiții, în anumite specii. Pe scurt, trebuie să apară. După cum tocmai am menționat, nu am nicio problemă cu această noțiune, evident. Însăși inima modelului **CBC** este înrădăcinată în presupunerea că punctul de apariție este prima apariție a organismelor unicelulare mobile. Aici vor trebui rezolvate probleme biochimice dificile, care probabil vor fi paralele (dacă nu identice) cu cele implicate în stabilirea a ceea ce determină apariția vieții din noroiul primordial, fără viață, care a precedat-o.

Dar acest efort de a justifica panpsihismul prin juxtapunere cu *emergentismul* nu oferă un răspuns satisfăcător. Nu face decât să dea cu piciorul (sau să urce?) pe drum și pare să aibă sens doar atunci când o serie de lucruri sunt omise în mod convenabil. Dacă mintea sau conștiința penetrează totul, cum a apărut această caracteristică **pan**-istă? De unde a apărut această chestie mentală care pătrunde miraculos? Singurul răspuns logic, non-misterios, este că trebuie să fie produsul unor proprietăți ale Big Bang-ului din care s-a născut universul nostru. Pe scurt, a „apărut” la acel moment în timpul cosmic. Deci, totul este de același fel. Dacă materia materială generează materia mentală, trebuie să existe un set de caracteristici, forme, funcții, structuri sau mecanisme prin care acest lucru s-a întâmplat prima dată. Acesta este punctul pe care panpsihicii precum Chalmers și Goff tind să evite să îl discute, deoarece versiunea lor a doctrinei se bazează pe prezumția că mentalul nu poate proveni din material, că mintea nu poate proveni din creier în niciun mod direct, fizic.

Argumentul lipsei dovezilor.

Poziția panpsihică este complet lipsită de date care să o susțină, ceea ce este, în acest moment, o problemă ireconciliabilă. O modalitate de a vedea cât de goală este poziția din punct de vedere empiric este să luăm notă de ceea ce le place panpsihicilor să vorbească: conștiință, minți, un univers pătruns în întregime de o conștiință fenomenală cosmică, lucruri care sunt pline de sensibilitate. De obicei, ei nu vorbesc despre lucruri inerte, pietre sau prăjitoare de pâine sau lună sau materie întunecată sau găuri negre. Iar atunci când o fac, susțin fără logică sau dovezi empirice că pietrele, prăjitoarele de pâine și materia întunecată au o esență. Desigur, nu există date care să susțină o astfel de afirmație. Dacă există ceva esențial în legătură cu o piatră, acesta este faptul că este o piatră. Nu există cunoaștere, nu există acțiune, nu există experiență. Nu există conștiință. Nu există nimic din ceea ce înseamnă să fii o piatră. Am argumentat deja că computerele nu știu nimic, nu au o sensibilitate ontologic coerentă, așa că nu este prea greu să negăm aceste stări fenomenologice pietrelor și prăjitoarelor de pâine.

Chalmers și alții au făcut eforturi neconvingătoare pentru a susține că obiecte precum termostatele au experiențe mentale. Dar nu i-am văzut, nici pe ei, nici pe altcineva, jucând jocul cu piatra sau cu prăjitorul de pâine. Pare corect să întrebăm: dacă această conștiință omniprezentă despre care panpsihicii susțin că pătrunde în orice, care sunt proprietățile sensibile ale acestor aparate de bucătărie moarte? Ce anume le face moarte, în timp ce lucrurile organice sunt vii? Dacă prăjitoarele de pâine sunt înzestrate cu acest elan vital, de ce nu îl manifestă? De ce compoziția lor fizico-chimică nu manifestă nicio sensibilitate? Ce se întâmplă atunci când o persoană moare? Creierul său acum disfuncțional capătă brusc o altă formă de conștiință? Întregul eșafodaj retoric este un sat Potemkin filosofic complicat. Pur și simplu nu are niciun sens - un aspect subliniat în mai multe rânduri de mulți, inclusiv, deloc surprinzător, de John Searle, care a declarat într-o

recenzie a cărții lui Christof Koch, *Consciousness: Confessions of a Romantic Reductionist* pentru The New York Review of Books:

Problema cu panpsihismul nu este că este fals; el nu se ridică la nivelul de a fi fals. Este strict vorbind lipsit de sens pentru că nu a fost dată nicio noțiune clară afirmației. Conștiința vine în unități, iar panpsihismul nu poate specifica unitățile.

După cum Searle arată foarte clar, dacă nu poți specifica unitățile, nu poți nici măcar să începi să colectezi date, iar fără suport empiric pentru o idee nebunească, nu are niciun rost să o iei în considerare.

Argumentul de pretutindeni.

Din punct de vedere metafizic, există o altă problemă, la fel de fundamentală, cu panpsihismul - și anume, presupunerea „peste tot". Dacă ceva este pretutindeni, ar putea la fel de bine să nu fie nicăieri. Existența este determinată de nonexistență. Când mă uit la o vază, îmi dau seama unde este pentru că îi simt contururile. Înțeleg unde este pentru că simt unde nu este. Limitele sunt esențiale; ele marchează existența. „A nu fi" este cheia către «a fi». Peștele va fi ultimul care va descoperi apa.[40]

Ca atare, conștiința presupusă, evazivă și nematerială pe care panpsihicii o presupun ca fiind prezentă în întreaga existență nu are nicio putere explicativă ca propunere științifică. Ea poate avea un element vag satisfăcător, deoarece atenuează nedumerirea care rezultă adesea din privirea prea lungă asupra problemei dificile, dar nu pare să poată juca un rol într-o știință serioasă a conștiinței. În treacăt fie spus, ar trebui să menționăm că există și o versiune mai blândă a panpsihismului, introdusă de Ernst Hackel și apărată de politologul William Clifford, care limitează domeniul de aplicare al conștiinței la organismele vii. Denumită uneori *panpsihism darwinist*, aceasta pornește, în esență, de la aceleași ipoteze ca modelul **CBC** - și anume, că toate organismele vii sunt înzestrate cu viață mentală, sensibilitate sau, dacă doriți, conștiință. Panpsihicii de acest tip ar fi mai mult decât confortabili cu **CBC**. L-ar îmbrățișa și m-ar primi la bord, unde am putea discuta dacă plantele sunt incluse - un subiect pe care îl vom aborda în capitolul 5.

În timp ce mă gândeam la motivul pentru care panpsihismul pare să fi răsărit ca Lazăr din fruntea atâtor cercetători contemporani din diverse domenii de studiu, m-am întrebat dacă nu asistăm la ceva asemănător cu ceea ce psihologul social canadian James Alcock a observat atunci când a examinat credința în paranormal.[41] El a observat că mulți, oameni de știință și profani deopotrivă, păreau să găsească lumea dură și neclintită a științei satisfăcătoare intelectual, dar neîmplinitoare spiritual. Parapsihologia era un loc natural către care să se îndrepte. Ea permitea îmbrățișarea enigmaticului, a transcendentului, fără a respinge știința empirică, deoarece se puteau face cercetări de laborator asupra unor lucruri precum percepția extrasenzorială și precogniția sau se puteau efectua studii de caz asupra unor persoane care susțineau că au fost răpite de extratereștri. Acesta a reprezentat un compromis frumos, având în vedere că satisface ambele pasiuni, ambele seturi de nevoi.

Nu am nicio idee dacă acest tip de dinamică funcționează în mințile panpsihicilor moderni, dar nu aș fi surprins dacă ar fi așa. După cum a remarcat filosoful Robert McCauley în volumul său tranșant din 2011, *Why Religion is Natural and Science is Not*, este ușor să fii religios; este greu să devii om de știință. Poate fi chiar mai greu să începi ca om religios și să treci la cel de-al doilea fără a purta vestigiile afective ale vieții tale anterioare.

Mai târziu vom analiza mai îndeaproape varietățile a ceea ce se numesc perspective filosofice misteriene - o acuzație de care panpsihicii nu pot scăpa. Aceste orientări obscurantiste

---

40 Există o lipsă de consens cu privire la originea acestei zicale asemănătoare unui proverb. Ea a fost atribuită uneori vorbitorilor Ewe din sudul Ghanei, psihologului Jerome Bruner, filosofului canadian Marshall McLuhan și unui budist tibetan anonim. Fără alte dovezi în afară de prejudecățile mele personale, îl susțin pe budist.

41 A se vedea cartea lui Alcock din 1981, Parapsihologia: Science or Magic (Știință sau magie), în care acest aspect este prezentat împreună cu o critică dură a teoriilor, metodelor de cercetare și analizelor de date din domeniul parapsihologiei.

sunt conduse de presupuneri particulare, nu cu privire la minte și conștiință, ci cu privire la ceea ce ele susțin că este golul de netrecut dintre creier și minte, credința lipsită de dovezi, dar larg răspândită, că funcțiile mentale, experiența fenomenală nu pot fi reduse la substanțe materiale, la creiere, structuri corticale, rețele neuronale sau mecanisme biologice moleculare. Întotdeauna există, susțin ei, acest decalaj explicativ între materie și minte.

Deși rațiunea iese adesea pe ușă atunci când sunt ridicate astfel de probleme, o replică rezonabilă în acest caz ar părea să fie: „Uitați, oameni buni, suntem absolut siguri că există principii biofizice fundamentale care acționează aici. Cunoaștem unele dintre ele; doar că nu le-am elaborat încă pe toate". Fizicienii încă nu știu cum să ajungă de la principiile mecanicii cuantice la o ceașcă de cafea, dar nimeni (sper) nu spune că „golul explicativ" al fizicienilor ar trebui să îi împingă (sau să ne împingă) spre moduri misterioase și antimaterialiste de a gândi despre materie.

Modelul **CBC** umple golul conștiinței sau, cel puțin, stabilește unde să căutăm pasajul, punctul biochimic al apariției. Repet: teza teoriei **CBC** este că cel mai probabil și cel mai logic moment pentru ca acest lucru să se fi întâmplat, pentru ca sensibilitatea să fi dansat pentru prima dată în fața reflectoarelor realității, este prima apariție a vieții. Acesta, susțin eu, este un moment emergent mai rezonabil și cu siguranță mai ușor de urmărit decât încercarea de a extrage sensibilitatea fie din momentul Big Bang-ului, fie, dacă doriți, din așezarea universului în forma sa actuală în urmă cu aproximativ șapte miliarde de ani - mai ales când sunteți blocați în încercarea de a vă da seama ce să faceți cu prăjitoarele de pâine.

Să revenim acum la cei care ar putea fi de acord cu celulele sensibile.

Michael Gazzaniga.

În noua sa carte (*Conscious Instinct.Unraveling the Mystery of How the Brain Makes the Mind*/Instinctul conștient.Dezlegarea misterului modului în care creierul creează mintea) Gazzaniga vede conștiința ca pe un instinct. Termenul de instinct în sine este în mod notoriu alunecos, dar el îl definește simplu ca fiind „ceva cu care un organism vine echipat". Având în vedere că mi-am stabilit deja poziția în ceea ce privește utilizarea termenilor tehnici în moduri care țin de psihologia populară, nu am nicio problemă cu acest lucru - deși alții, bănuiesc, ar putea avea.

La fel ca majoritatea neuroștiințelor care se îndreaptă spre puzzle-ul conștiinței, adevărata țintă a lui Gazzaniga este conștiința umană care, concluzionează el, ar trebui privită cel mai bine, nu ca o construcție mentală singulară, ci ca funcțiile complexe interconectate ale unei serii de module, fiecare dintre acestea evoluând pentru a avea propria structură, formă și funcție particulară. În virtutea poziției sale evoluționiste mai largi, pare probabil că va fi confortabil cu modelul **CBC**. După cum notează el,

Ființele vii au o organizare care permite existența vieții și, în cele din urmă, a conștiinței, chiar dacă sunt făcute din aceleași materiale ca și lumea naturală nevie care le înconjoară. Iar instinctele învăluie organisme de la bacterii la oameni. (Subliniere adăugată.)

Gazzaniga observă, de asemenea, că dintre numeroasele „lacune explicative" pe care filosofii și neuroștiințele le-au identificat (și cu care s-au luptat), cel mai important este cel dintre organizarea moleculelor care nu este vie și cea care este - un aspect important și adesea neglijat.

Lynn Margulis.

Într-o lucrare din 2001 intitulată provocator „*Celula conștientă*”, controversatul[42] microbiolog Lynn Margulis a precizat clar care este poziția sa în această privință. În rezumat, ea se referă la o „*minte microbiană*” și, deși lucrarea se axa în primul rând pe simbiogeneză și pe rolul acesteia în crearea eucariotelor din bacterii care înainte trăiau liber, ea a fost lipsită de ambiguitate în credința sa în sensibilitatea primelor forme organice:

Înregistrarea fosilelor de eucariote începe acum aproximativ 2 000 de milioane de ani. Primele animale marine cu coajă tare au apărut acum 541 de milioane de ani. Aceasta datează evoluția părților dure. Creierele apar mai târziu, dar conștiința, conștientizarea mediului înconjurător, începe cu începutul vieții înseși (sublinierile sunt adăugate).

Acea „celulă conștientă” din titlu pare aproape o idee ulterioară, o replică de aruncat al cărei adevăr părea, în mintea ei, prea evident pentru a necesita o elaborare.

Peter Godfrey-Smith.

În recenta sa carte *Other Minds: The Octopus, The Sea and the Deep Origins of Consciousness*, biofilosoful Peter Godfrey-Smith a dat, de asemenea, din cap posibilitatea existenței minții la speciile unicelulare:

[Am] putea concluziona că toate ființele vii au un minim de experiență subiectivă - o opinie pe care nu o consider nebunească, dar care ar necesita cu siguranță multă apărare .....

Godfrey-Smith se simte destul de confortabil în ceea ce privește înzestrarea caracatițelor și a altor cefalopode precum sepia cu minți, dar ezită să ia în considerare speciile mai puțin dezvoltate. Se pare că nu se simte în largul său atunci când vorbește despre sensibilitatea autentică a organismelor cărora le lipsește un sistem nervos de orice fel. Mai târziu, vom analiza din nou poziția lui Godfrey-Smith și a altora care consideră că conștiința se găsește în multe specii, altele decât oamenii. Principalele dispute între astfel de gânditori se referă la aspecte precum: Ce fel de sistem este necesar pentru a exista stări mentale autentice? Cât de complex trebuie să fie sistemul? Ce tipuri de funcții ar trebui să fie capabil să exprime? Bănuiesc că cercetători precum Godfrey-Smith ar trebui (ar putea? ar trebui?) să fie dispuși să accepte principiile cadrului **CBC** dacă consideră că apărarea mea a acestuia (care este prezentată în capitolul 4) este convingătoare.

Maxim Sheets-Johnstone.

În cartea sa fundamentală *The Primacy of Movement*, Sheets-Johnstone a dezvoltat argumentul conform căruia conștiința nu poate fi înțeleasă, sau chiar contemplată într-o manieră sensibilă, fără a aprecia că mințile nu apar pur și simplu din creiere, ci că toate experiențele, toate sentimentele, apar în organisme care se mișcă, navighează în medii complexe, în continuă schimbare - și că natura locomoției lor și modul în care sunt formate sunt dictate de corpurile lor, de formele lor, de forma (formele) lor de mișcare și de capacitățile lor senzori-motorii. Vom avea mai multe de spus despre modelul lui Sheets-Johnstone, deoarece acesta este foarte deschis la ideea că viața mentală este o caracteristică inerentă a tuturor formelor organice. O includ aici pentru că afirmă, fără ambiguitate, că forme primitive de conștiință trebuie să fi fost prezente în orice specie

42 Margulis a fost o adevărată excentrică științifică. De-a lungul vieții sale îndelungate, ea a susținut o serie de poziții care pot fi mai bine numite „aberante”. Unele au fost introduse în cele din urmă în modelul standard (poziția sa privind simbiogeneza), altele au rămas marginale (ipoteza Gaia, care consideră Pământul ca o entitate vie). În ciuda criticilor viguroase la adresa activității sale, ea și-a învins în cele din urmă (majoritatea) detractorilor. Președintele Clinton i-a acordat Medalia Națională pentru Știință, iar Societatea Linneană i-a acordat prestigioasa Medalie Darwin-Wallace.

mobilă, inclusiv în organismele unicelulare. Într-o lucrare recentă care comentează articolul meu țintă din *Animal Sentience*, ea a aprobat (mai mult sau mai puțin) **CBC** ca fiind o teorie viabilă.

Frantisek Baluska.

Un biolog al celulelor vegetale de la Universitatea din Bonn, Baluska s-a concentrat asupra funcțiilor de comunicare și semnalizare la plante și a făcut unele descoperiri fascinante în ceea ce privește funcțiile rădăcinilor care, a observat el, funcționează în moduri analoge cu sistemele nervoase ale animalelor. Într-un interviu recent cu jurnalista științifică Suzan Mazur, Baluska a declarat că a fost inspirat de lucrarea lui Lynn Margulis pe care am menționat-o mai sus și a comentat,

Viața începe atunci când un sistem începe să fie un agent independent conștient de sine. Acesta se străduiește să își protejeze independența. Sistemul are imagini de feedback care îi oferă în mod constant informații pe care le-a preluat din mediul în care funcționează. Este conștient de sine și capabil să își păstreze integritatea.

Vom analiza în detaliu activitatea lui Baluska și a colaboratorilor săi în capitolul 5, când ne vom confrunta cu întrebarea dacă flora este sensibilă. Pentru a anticipa, el a susținut că funcțiile surprinzătoare ale plantelor sunt de așa natură încât acestea ... pot fi calificate ca fiind cognitive". Plantele sunt eucariote, iar lucrarea lui Baluska nu se referă la procariote, dar bănuiesc că va considera că abordarea CBC este rezonabilă. Citatul de mai sus, care ne invită să echivalăm viața și conștiința de sine, se generalizează cu ușurință la speciile unicelulare.[43]

Dennis Bray.

Bray, microbiologul de la Cambridge, a fost cât pe ce să îmbrățișeze ceva de genul **CBC** fără să muște glonțul metaforic. În cartea sa din 2009, *Wetware*, el scrie poetic,

Este posibil ca o amibă să posede, într-o formă extrem de redusă și primitivă, unele dintre mecanismele care mediază sentimentul de mediu și de sine la oameni? Ce cunoștințe are o celulă despre ea însăși? Este complet întuneric acolo, un spațiu negru plin de mașinării moleculare oarbe? Sau s-ar putea ca acea suspensie apoasă să conțină o scânteie de emoție, un prototip de sensibilitate?

Atunci când a fost forțat să răspundă la propria întrebare, el a refuzat să facă acest pas.

În opinia mea, o celulă individuală este un sistem care posedă ingredientele de bază ale vieții, dar căruia îi lipsește sensibilitatea. Este un robot făcut din materiale biologice.

Într-o discuție privată, Bray m-a informat că, de fapt, s-a răzgândit și că astăzi ar răspunde afirmativ la propria sa întrebare. Dar, indiferent de poziția actuală a lui Bray în ceea ce privește chestiunea sensibilității procariotelor, *Wetware* oferă o prezentare superbă a funcției și formei microbiologice, folosind, după cum sugerează titlul, inteligența artificială și robotica bazate pe computer ca dispozitive retorice. M-am bazat pe materialul din volumul lui Bray atunci când am dezvoltat (cât de bine poate acest psiholog cognitiv cu probleme microbiologice) funcțiile biochimice-mecanice care stau la baza acelei creaturi unicelulare despre care amândoi credem că are o minte.

43 Baluska și cu mine am prezentat recent lucrări și am participat la un forum deschis în cadrul Institutului de la UQAM menționat anterior. El a confirmat că consideră că principiul de bază al **CBC** este valabil și aproape sigur adevărat.

Alţii, inclusiv David Edelman, Roger Penrose şi Stuart Hameroff, şi chiar Daniel Dennett.

Biologul David Edelman, împreună cu psihologul Bernie Baars şi neurologul Anil Seth, a desfăşurat o amplă activitate privind conştiinţa nevertebratelor, care sugerează cu siguranţă că s-a gândit să îşi extindă descoperirile până la organismele unicelulare (pentru cei interesaţi de detalii, accesaţi Edelman. Baars. & Seth. 2005 şi Edelman & Seth, 2009). Abordarea fizicianului Roger Penrose şi a anestezistului Stuart Hameroff, care este discutată în detaliu mai târziu, ar trebui să îi lase deschişi la posibilitatea de a îmbrăţişa modelul **CBC**. Abordarea lor presupune că funcţionarea internă a microtubulilor din celulele creierului duce la conştiinţă, deşi abordarea lor se bazează pe principiile mecanicii cuantice şi este destul de diferită de cea dezvoltată aici (şi de cea a tuturor celorlalţi, de altfel). Daniel Dennett, de asemenea, se apropie, dar în cele din urmă se îndepărtează. În ultima sa lucrare, *From Bacteria to Bach and Back*, are o secţiune în care compară „minţile lifturilor" (adică programele de calculator care gestionează lifturile) cu „minţile înălţate" sofisticate (adică ale noastre). Aici inserează această notă de subsol:

Recunosc că este posibil ca ele [bacteriile] să [aibă „minte"]. Tratez această posibilitate ca fiind neglijabilă, un risc strategic non-zero pe care sunt pregătit să mi-l asum. (Sublinierea este în original.)

Dennett pare să fie pregătit să acorde propunerii conform căreia bacteriile au minţi o probabilitate diferită de zero de a fi adevărată. De asemenea, dacă este dispus să recunoască faptul că, în cazul în care acestea experimentează forme elementare de sensibilitate, acestea sunt stări mentale valide din punct de vedere ontologic şi nu „iluzii ale utilizatorului", i-aş ura bun venit la bord. Poziţia lui Dennett, pe care o vom analiza mai îndeaproape mai târziu, abordează conştiinţa umană asemănând-o cu o „iluzie a utilizatorului", în care funcţionarea internă naturală a creierului nostru este interpretată în mod eronat ca fiind acţiunile de control ale unei conştiinţe ontologice.

Pentru Dennett, conştiinţa este epifenomenală, resturile mentale ale funcţiilor cerebrale lipsite de orice putere cauzală. Este greu de imaginat că o minte atât de primitivă precum cea care operează în *Escherichia* coli ar putea genera o iluzie de utilizator a conştiinţei, deoarece, în modul de gândire al lui Dennett, ar trebui să aibă capacitatea de a reacţiona mental la propriile funcţii implicite. Din aceste motive, **CBC** poate fi văzut şi ca un antidot la neo-comportamentalismul rezidual al lui Dennett.

Sacks, Damasio, Cook, Cavalho şi Hunt.

În 2015, aceşti cinci distinşi neurocercetători au scris o scrisoare deschisă remarcabilă către neurocercetătorul Christof Koch[44], în care îi propuneau că cercetarea caracteristicilor esenţiale ale conştiinţei ar fi avantajată de explorarea aprofundată a caracteristicilor microbiologice specifice ale speciilor unicelulare. Scrisoarea, care a apărut în *Scientific American*, este fascinantă, la fel ca şi răspunsul lui Koch la aceasta. În scrisoare, Sacks et al. recomandă o abordare a microbiologiei foarte asemănătoare cu **CBC**, care,

*implică căutarea proprietăţilor care stau la baza conştiinţei până la nivelul protozoarelor ... pentru a identifica mecanismele fundamentale la nivel celular care, atunci când sunt extinse în sisteme nervoase complexe, dau naştere proprietăţilor care sunt denumite în mod obişnu*it „minte".

44 Koch, a cărui poziţie cu privire la conştiinţă am abordat-o anterior (şi o vom analiza mai îndeaproape mai târziu), este preşedinte şi director ştiinţific la Allen Institute for Brain Science din Seattle. De asemenea, acesta face parte din consiliul de consilieri al Scientific American Mind, ceea ce, probabil, este motivul pentru care Cook et al. i-au trimis scrisoarea.

Ei au continuat să identifice unele proprietăți ale celulelor vii despre care au speculat că sunt candidați rezonabili pentru apariția sensibilității, concentrându-se în principal pe ceea ce au numit iritabilitate sau exprimarea unei hipersensibilități caracteristice tuturor celulelor. Această reacție, susțin ei,

este probabil captată de atacul brusc de ioni încărcați pozitiv (cationi) în citoplasma alcalină - definiția însăși a excitabilității membranei - [și] este fenomenul-cheie implicat în „conștientizarea" de către o celulă a mediului său („sensibilitate").

Mai mult decât atât, ei susțin că acest mecanism, care este cauza biochimică de bază a excitabilității membranei, poate fi de fapt văzut ca o soluție la enigma pe care am urmărit-o, și anume

„problemele dificile" ale sensibilității, conștiinței și, în cele din urmă, conștiinței de sine a primatelor [care] încep cu răspunsul celulelor excitabile la stimuli externi care amenință să perturbe homeostazia celulară.

Pe scurt, ei au adoptat o poziție care împărtășește caracteristicile de bază ale cadrului **CBC**. După cum vom vedea mai târziu în capitolul 4, unde modelul **CBC** este dezvoltat mai detaliat, sunt agnostic în ceea ce privește rolul critic al excitabilității membranei celulare în apariția sensibilității - deși sunt cu siguranță de acord cu acest lucru în cazul în care cercetarea empirică va demonstra că Sacks et al. au avut dreptate. Dar ceea ce mi s-a părut la fel de intrigant ca perspectiva lor a fost răspunsul lui Christof Koch. A fost clasic. Koch, unul dintre cei mai importanți neurocercetători ai generației sale și autor al unei abordări intrigante, dar aproape în întregime orientată spre *H. sapiens*, a conștiinței (*Consciousness: Confessions of a Romantic Reductionist*). Koch a lăudat autorii, a recunoscut poziția prezentată de aceștia, a recunoscut legitimitatea argumentelor microbiologice prezentate, apoi a plecat, distanțându-se de propunerea principală. El a observat că, deși influxul de cationi este esențial pentru excitabilitatea celulară și că mecanismele de permeabilitate a membranelor sunt caracteristici critice, este, în plus, adevărat că

de asemenea, nu am fi conștienți fără o inimă care bate, fără un flux sanguin oxigenat adecvat, fără un cocktail adecvat de substanțe chimice neuromodulatoare în diferite regiuni ale creierului, cum ar fi sistemul cortico-talamic și așa mai departe. Toate acestea sunt condiții de bază care permit apariția conștiinței, dar sunt insuficiente în sine. Este nevoie de mai mult.

Acest lucru este complet adevărat și total irelevant. Sacks și colegii săi prezentau un argument pe care Koch, se pare, fie l-a omis, fie l-a deviat în mod conștient. Niciunul dintre semnatarii scrisorii nu ar contesta afirmațiile lui Koch, dar toți le-ar considera o pistă falsă[45] - la fel ca mine. Nu conștiința primatelor (sau cefalopodelor sau păsărilor) îi preocupa. Ei susțineau că modul de a desluși în cele din urmă relațiile cauzale dintre toate organele și operațiunile neuromodulatoare care joacă roluri semnificative în conștiința umană este de a începe, nu cu genul *Homo*, ci cu protozoarele. Koch, care a revenit la o relatare centrată pe primate doar ușor deghizată, îi deviază pe Sacks și asociații săi, observând că ceea ce va fi necesar în cele din urmă este,

o explicație principială, analitică, prescriptivă, testabilă empiric și utilă clinic a modului în care <u>materia foarte organizată și excitabilă</u> susține faptul central al existenței noastre - experiența subiectivă, [în care]... excitabilitatea membranară va juca un rol aici, dar ca unul dintre mulți alți factori și considerații. (Subliniere adăugată.)

45 Un termen care datează de la tehnica de a târî un ardei bine afumat și înțepător în calea unui animal pentru a perturba mirosul și pentru a trimite urmăritorii pe o vânătoare de gâște sălbatice, amestecând metaforele aici. Ca dispozitiv retoric, poate fi folosit fie în mod deliberat (ca în multe dezbateri filosofice), fie inconștient, așa cum cred că a făcut Koch aici.

Koch își încheie răspunsul cu o analogie interesantă. El susține că încercarea de a înțelege conștiința (umană) pornind de la mecanismele de transport membranar ar fi ca și cum am încerca să înțelegem internetul printr-o descriere completă a fizicii tranzistorilor. Aceasta este o analiză interesantă, dar ratează, încă o dată, esența a ceea ce Sacks et al. propuneau. Deși nu au abordat nicio problemă de filozofie a științei în scrisoarea lor (a fost, la urma urmei, o misivă relativ scurtă), poziția pe care au adoptat-o este, de fapt, una viguros reducționistă. Ei prezentau propunerea implicită conform căreia conștiința umană și sensibilitatea oricărei alte specii ar putea foarte bine să fie reduse la rolurile cauzale pe care le joacă celulele vii individuale - la fel cum internetul poate fi redus la rolurile cauzale pe care le au cipurile de siliciu și tranzistorii. Ambele ar fi complexe și complicate, dar, în principiu, pot fi realizate.

Este important să apreciem ce a făcut și ce nu a făcut Koch în răspunsul său. El a recunoscut că funcțiile biochimice identificate de Sacks, Damasio, Cook, Cavalho și Hunt sunt elemente critice în definirea vieții, că sunt componente esențiale ale capacității unui organism unicelular de a simți ceva, că fac parte din capacitatea diverselor organisme de a reacționa la input-uri și de a face ajustările comportamentale adecvate - dar (întotdeauna există acel „dar” atunci când subiectul este conștiința), în opinia sa, „este nevoie de mai mult”. Desigur, motivul este că Koch, la fel ca mulți alții care abordează această problemă, prezintă acea ciudățenie cognitivă caracteristică, acea dificultate singulară de a disocia noțiunea de minte sau conștiință de instanțierea acesteia la oameni. Prin urmare, aceste alte caracteristici experiențiale pe care Sacks et al. le identifică ar putea avea ceva de-a face cu „sensibilitatea”, dar sunt insuficiente pentru a explica conștiința (umană). Mai târziu, după ce vom analiza modul în care alți cercetători s-au alăturat problemei dificile, vom analiza mai în profunzime de ce atât de mulți neuroștiințifici și filosofi sunt de aceeași părere ca Koch.

Rezumat

O scurtă trecere în revistă a stadiului în care ne aflăm până în prezent:

- Funcționalismul independent de hardware nu poate funcționa. A deveni anorganic este o fundătură conceptuală, științifică și filosofică. Conștiința, sensibilitatea, subiectivitatea sunt proprietăți ale formelor organice. Problema dificilă nu va fi rezolvată prin scrierea unui cod.
- Un număr semnificativ de persoane, de la scriitori de *science fiction* la microbiologi, filosofi, matematicieni, neuroștiințifici, etologi și psihologi cognitivi au susținut diverse teorii care se bazau pe principii biologice evolutive care ar trebui, ar putea, ar fi posibil sau aproape au presupus că subiectivitatea a apărut pe scenă odată cu primele licăriri de viață pe această planetă.
- Cu toate acestea, niciunul nu a apucat inelul de alamă în sensul recunoașterii faptului că implicațiile unui model precum **CBC** merg dincolo de concluziile privind conștiința de sine unicelulară. Odată ce se apreciază că organismele unicelulare au minte și subiectivitate, o serie de poziții empirice și filosofice capătă un aspect foarte diferit. Până în prezent, aceste implicații nu au fost recunoscute, cu atât mai puțin urmărite.

În capitolul următor vom analiza o varietate de modele „biocentrice” care au fost propuse și vom examina de ce nici acestea nu au reușit să formuleze un model coerent al conștiinței. După cum va deveni evident, deși există dificultăți empirice, conceptuale și teoretice enorme cu care se confruntă oricine se aventurează în acest domeniu epistemic, adevăratul vinovat este simplu și l-am depistat de timpuriu: acea fascinație excesivă față de propria noastră specie și îmbrățișarea narcisistă a caracterului special al conștiinței umane.

Această poziție, odată asumată, limitează destul de brusc gama de abordări științifice care vă sunt deschise. De fapt, nu vă rămân decât două. Fie încerci să înțelegi cum creierul uman creează mintea umană, fie cauți instanțieri ale conștiinței asemănătoare celei umane la alte specii mai puțin complexe și încerci să determini ce împărtășesc acestea (fie structural, fie comportamental) cu oamenii. După cum vom vedea, nicio strategie nu a funcționat. Problema dificilă rămâne nerezolvată. Lacuna explicativă seamănă mai degrabă cu o prăpastie de netrecut. Aceleași probleme filosofice persistă. Va fi util să înțelegem motivele lipsei de succes, așa că haideți să aruncăm o privire asupra acestor abordări și să vedem unde se află dificultățile.

3. Viziuni biocentrice asupra originii conștiinței: Căutarea soluției la o problemă dificilă

Așa cum am subliniat în capitolul 1, eșecul funcționaliștilor de a susține ideea unei conștiințe bazate pe calculatoare, independentă de hardware, a deschis calea pentru cealaltă linie de cercetare, abordarea biocentrică - deși trebuie să recunosc cu părere de rău că eșecul computaționaliștilor a fost un pic dezamăgitor. Dacă ar fi fost pe drumul cel bun și dacă simțirea autentică, conștiința și toate funcțiile pe care creierele ca al nostru le pot îndeplini ar fi fost reductibile la linii de cod, soluția la problema dificilă ar fi fost transformată într-una foarte ușor de gestionat - compilarea (multor) softuri și firmware-uri sofisticate.

Dar, întrucât majoritatea celor din domeniu recunosc în prezent că a capta conștiința autentică, subiectivitatea ontologică, într-o inteligență artificială (AI) a fost un vis irealizabil, accentul a fost pus pe științele neurobiologice. În ultimele câteva decenii, proiectele din domeniul neuroștiințelor cognitive au scos la iveală o serie de legături interesante între anumite funcții, procese, abilități și seturi de competențe și centrele, căile și rețelele neuronale care le modulează și le controlează. Biologii evoluționiști au identificat structuri neuronale în sistemele nervoase ale unei varietăți de alte specii care sugerează rădăcinile filogenetice ale multor funcții cognitive despre care se presupunea odată că sunt privilegiul exclusiv al creierului uman. Cercetarea este fascinantă, iar eu sunt, și am fost încă de când am intrat în școala doctorală, un susținător entuziast al acestei lucrări. Dar nu faceți nicio greșeală. Este o știință bună și solidă, dar nimic din toate acestea nu ne-a adus mai aproape de rezolvarea problemei dificile. Și, în mod interesant, niciuna dintre perspectivele sau abordările adoptate de numeroșii cercetători care lucrează în domeniu nu a arătat vreun indiciu de a deveni un cadru general acceptat în care să se lucreze. După cum au remarcat mulți, știința conștiinței se află într-un stadiu „pre-paradigmatic” kuhnian, cu o varietate de modele reciproc incompatibile care concurează între ele. Să aruncăm o privire rapidă asupra unora dintre cele mai influente abordări.

Teorii generale privind originile conștiinței

Au existat mai multe modele elaborate pe larg care încearcă să identifice sursa conștiinței și/sau să caracterizeze modul în care diferite proprietăți ale conștiinței sunt instanțiate în creier. Unele sunt prezentate ca soluții candidate la problema dificilă, altele pur și simplu ca ipoteze științifice cu privire la funcțiile cerebrale și rolul lor în sensibilitate. În aceste din urmă cazuri, filosofii ar putea considera cercetarea ca fiind relevantă pentru problemă, dar, de obicei, susținătorii încearcă doar să facă știință bună. Cu toate acestea, aproape toate pornesc de la cele aproximativ un kilogram de neuroni, glia, neurotransmițători, structuri de suport și apă care alcătuiesc creierul uman. Practic, toți presupun, într-un fel sau altul, că există centre neuronale identificabile, module, căi, mecanisme de procesare corticale și/sau subcorticale, funcții inter și intracelulare și legi biomecanice și/sau biofizice de bază cunoscute care dictează modul în care acest organ face ceea ce face.

Vom arunca o privire atentă (și critică) asupra a cinci dintre cele mai ambițioase, influente și creative eforturi, în special: modelul Penrose-Hameroff, care se bazează pe ipoteza unor efecte cuantice în celulele nervoase; teoria neurală darwiniană a lui Gerald Edelman, care adaptează ideile sale anterioare din imunologie și le transpune în simțire; abordarea lui Daniel Dennett, care se bazează pe o analiză atentă a rolului selecției naturale darwiniene; modelul „post-cartezian” al lui Antonio Damasio, care pune accentul pe afect și emoție; și teoria Tononi-Koch, bazată pe teoria informației integrative (IIT). Vom face, de asemenea, o scanare rapidă a altor câteva care au avut un impact asupra domeniului, deși acestea aruncă o plasă conceptuală mai mică. Nu intenționez să transform această secțiune într-o analiză de tip manual. Vreau doar să transmit o idee despre

direcţiile pe care le-au luat neuroştiinţele şi filozofii în efortul de a se ocupa de această problemă extrem de interesantă şi supărătoare.

Penrose şi Hameroff.

Unul dintre primele modele, şi probabil cel mai inovator model, a fost cel introdus de celebrul matematician şi fizician britanic Sir Roger Penrose şi dezvoltat în detaliu în colaborarea sa continuă cu anestezistul american Stuart Hameroff (a se vedea Hameroff & Penrose, 1996. 2014; Penrose, 1989; Penrose & Hameroff. 1995).

Penrose a abordat problema biofilosofică a conştiinţei dintr-un mediu academic îndepărtat, fizica matematică. Interesul său a fost stârnit de posibilitatea ca efectele cuantice să se producă la nivelul neuronilor din creier şi că aceste operaţiuni nealgoritmice ar putea duce la manifestarea experienţei conştiente. Cartea lui Penrose din 1989, *The Emperor's New Mind* (Noua minte a împăratului), a fost o călătorie sclipitoare şi extinsă în jurul ezoteriei ştiinţei. Am luat-o pentru prima dată după ce am văzut un anunţ care anunţa că undeva în cele câteva sute de pagini Penrose va reuşi să discute:

numere complexe, maşini Turing, teoria complexităţii, mecanica cuantică, sisteme formale, Godel indecizia, spaţii de fază, spaţii Hilbert, găuri negre, găuri albe, radiaţia Hawking, entropie, quasicristale [şi, desigur] structura creierului.

Am băgat-o sub braţ şi m-am îndreptat spre casierie. Şi, într-adevăr, era un tur de forţă care viza în întregime, în cele din urmă, concluzia că operaţiunile fundamentale ale neuronilor din creierul uman nu urmau legi obişnuite, ci erau supuse lumii ciudate şi minunate a fizicii cuantice - o noţiune extrem de puţin plauzibilă, dar atât de îndrăzneaţă încât nu puteai să o arunci pur şi simplu în cel mai apropiat tomberon. Atenţia era obligatorie.

Scopul declarat al lui Penrose a fost să argumenteze împotriva posibilităţii *inteligenţei artificiale puternice*, o poziţie pe care am susţinut-o din toată inima. Dar Penrose a mers pe o cale destul de diferită de cea pe care am urmat-o în capitolul 1. Argumentele sale au pornit de la presupunerea că funcţiile creierului nu pot fi calculate de dispozitive digitale deoarece, spre deosebire de aceste sisteme bazate pe siliciu, procesele neuronale care stau la baza acestora se bazează pe principii de mecanică cuantică, implicând posibil procesele încă necunoscute ale gravitaţiei cuantice. Sincer să fiu, când am citit această carte, primul meu gând a fost că, dacă Penrose nu ar fi fost unul dintre cei mai respectaţi fizicieni matematicieni din Regatul Unit (în 1988 a împărţit prestigiosul Premiu Wolf cu Steven Hawking, în 1989 a primit Medalia Dirac şi a fost făcut cavaler în 1994 pentru serviciile şi contribuţiile sale la ştiinţă), cartea ar fi ajuns rapid pe masa „*alege tu, 5 p*” din librăriile pitoreşti de cartier. În schimb, a devenit un bestseller, a generat o dezbatere animată şi a atras interesul unei serii de gânditori, inclusiv al lui Stuart Hameroff care, cu doi ani mai devreme, scrisese *Ultimate Computing*, în care speculase că procesele complexe din microtubulii celulelor creierului ar putea fi sursa conştiinţei şi a conştiinţei.

Împreună, ei au dezvoltat ceea ce a devenit teoria **Orch OR** („*Orchestrated Objective Reduction*” - „Reducere obiectivă orchestrată”), bazată pe două argumente: conştiinţa trebuie să implice elemente necomputaţionale şi astfel de stări interne pot fi surprinse de o coerenţă cuantică propusă în microtubulele găsite în celulele nervoase cerebrale. Teoria, poate deloc surprinzător, a devenit ţinta unor critici viguroase din mai multe părţi. În timp, s-au putut distinge patru tipuri de atacuri:

- *cele care susţineau că modelul nu a reuşit să stabilească un contact coerent cu neuroştiinţele de bază* (Koch & Hepp, 2006),
- *cei care au pus la îndoială faptul că presupusele principii fizice ar putea funcţiona efectiv în medii organice precum creierul uman* (Tegmark, 2000),

- *cele care au subliniat lipsa de sprijin empiric pentru ipotezele privind efectele cuantice în celule, în special în celulele neuronale din creierul uman* (Litt et al. 2006) şi, în sfârşit,
- *cele care au suspectat că aceste efecte cuantice pur şi simplu nu sunt posibile* (Georgiev, 2007).

Deşi Hameroff şi Penrose au continuat să dezvolte şi să apere teoria generală (a se vedea lucrarea lor din 2014 „*Consciousness in the universe*"), aceasta este privită astăzi ca o soluţie aberantă de către majoritatea microbiologilor şi neuroştiinţelor cognitive. Eu nu o consider un cadru viabil nici teoretic, nici empiric - dar am vrut să o ascult. Am o slăbiciune pentru teoriile radicale care ies din curentul epistemic actual.

Gerald Edelman.

Edelman, laureat al Premiului Nobel pentru Fiziologie sau Medicină pentru lucrările sale asupra sistemului imunitar, s-a aplecat, mai târziu în carieră, asupra problemei noastre mereu seducătoare - cum poate un creier să creeze o experienţă conştientă? Răspunsul său a fost încorporat în teoria sa a darwinismului neuronal care, în mod curios, se bazează pe aceleaşi principii pe care le-a descoperit în cercetările sale Nobel privind modul în care funcţionează limfocitele. Înainte de Edelman (şi Rodney Porter, biologul britanic cu care a împărţit premiul), modelul standard presupunea că există un singur tip de limfocite şi că acestea îşi modifică funcţiile anticorpilor pentru a se adapta la proprietăţile invadatorilor.

Cu toate acestea, Edelman şi Porter au descoperit că există o gamă largă de tipuri diferite de limfocite în organism, iar modul lor de funcţionare diferă. Atunci când apare un virus invadator sau un alt agent patogen, toate se grăbesc să apere organismul împotriva acestuia, concurând literalmente între ele pentru a-şi îndeplini sarcina. Dintre numeroasele varietăţi de anticorpi, cei care sunt cei mai eficienţi în distrugerea antigenelor invadatorului obţin un avantaj reproductiv. Acestea sunt, în sens darwinist, selectate şi se înmulţesc. Ei expediază inamicul, se stabilesc şi îşi menţin prezenţa continuă în sistem, pregătiţi pentru o ripostă rapidă în cazul în care aceiaşi antigeni îşi fac din nou apariţia. Acest mecanism, desigur, este modul în care funcţionează imunizarea şi motivul pentru care ar trebui să vă faceţi vaccinul antigripal în fiecare an şi să vă asiguraţi că copiii dumneavoastră urmează programul standard de imunizare recomandat.

Începând cu începutul anilor 1980, Edelman a propus o imagine a modului în care conştiinţa ar putea apărea din funcţiile creierului, folosind un analog al principiilor imunologice pe care le descoperise. El a presupus că grupurile de neuroni sunt selectate în acelaşi mod în care au fost selectate acele varietăţi de limfocite. Povestea spusă de Edelman începe la naştere, când un copil ajunge pe scenă echipat cu aproximativ 100 de miliarde de neuroni şi câteva instincte de bază sau tendinţe naturale integrate, cum ar fi atenţia selectivă, impulsurile esenţiale, sentimentele pentru stimuli negativi şi pozitivi şi altele asemenea. Edelman a presupus că unele dintre acestea sunt controlate de un sistem neuronal segregat, orientat spre procese interne, hedonice de reglare, în timp ce altele se concentrează pe codificarea informaţiilor externe privind starea lumii înconjurătoare. S-a presupus că aceste sisteme neurologice permit sugarului să distingă sinele de non-eu, ceea ce reprezintă punctul de plecare al experienţei „Cum este să fiu eu".

Odată cu creşterea, apar modele neuronale subiacente în creier, pe măsură ce creierul răspunde la stimulii primiţi. În modelul lui Edelman, fiecare nou stimul de intrare (un sunet nou, un percept vizual) este tratat într-o manieră analogă cu modul în care un invadator purtător de antigene este tratat de sistemul imunitar. Mai multe grupuri de neuroni răspund la acesta. Cei care „se potrivesc", care sunt configuraţi într-un mod care le permite să răspundă mai adecvat sau mai eficient la input sunt, în mod darwinist, selectaţi. Dar, din moment ce Edelman vorbeşte despre structurile neuronale din creier, nu despre limfocite, „selecţie" înseamnă aici că conexiunile lor sinaptice sunt consolidate în detrimentul altora care nu sunt. Cu intrări repetate, aceste conexiuni neuronale interconectate tind să formeze grupuri neuronale care devin stabile şi funcţionează ca

material neuronal de bază pentru o categorie de obiecte[46], în timp ce altele se ofilesc şi sunt eliminate.

Grupurile de neuroni care sunt păstrate interacţionează între ele şi, ca răspuns la aportul continuu, formează ceea ce Edelman a numit „hărţi". Hărţile sunt teoretizate ca fiind colecţii mari, sub formă de foi, de milioane de neuroni şi conexiunile lor sinaptice care corespund unor puncte echivalente dintr-un receptor (de exemplu, membrana bazilară din urechea internă, pielea, retina). Se crede că hărţile devin legate între ele şi se presupune că semnalele neuronale se deplasează în mod regulat înainte şi înapoi între ele într-un proces masiv paralel pe care Edelman l-a numit „*reentry/reintrare*". În opinia sa, conştiinţa nu ar trebui să fie reificată, deoarece nu este un lucru sau un obiect. Mai degrabă, ar trebui privită ca un proces care nu se regăseşte în niciun grup de celule sau căi. Conştiinţa, a presupus el, este distribuită pe scară largă în întregul creier şi rezultă din interacţiunile dintre amintirile experienţelor anterioare şi intrările prezente în mapările globale.

Viziunea lui Edelman s-a bucurat de un oarecare sprijin din partea altor oameni de ştiinţă, în special din partea regretatului neurolog şi eseist Oliver Sacks, precum şi a filosofului John Searle. Cu toate acestea, ea a atras mult mai multe critici, multe dintre acestea concentrându-se pe lipsa de claritate a expunerii. Chiar şi Searle a remarcat, într-o prezentare generală în mare parte pozitivă, că, deşi Edelman a abordat multe moduri în care sistemele cerebrale ar putea provoca anumite tipuri de stări de sensibilitate, nu este uşor să găseşti detalii despre modul în care apare de fapt această conştiinţă. Există o mulţime de detalii, multe dintre ele speculative, despre modul în care funcţionează creierul, dar nimic care, după cum a spus scriitorul de ştiinţă John McCrone într-o recenzie din 2004 a cărţii lui Edelman, *Wider than the Sky: The Phenomenal Gift of Consciousness*,

Unde Edelman eşuează cu adevărat este în încercarea de a trece de ultimul obstacol filosofic, punctul în care scepticii spun că este posibil să fi explicat funcţionarea creierului şi comportamentul inteligent care rezultă, dar acest lucru încă nu spune de ce activitatea creierului trebuie să se simtă ca ceva subiectiv, de ce trebuie să existe acea stranie strălucire interioară a experienţei.

Desigur, nu ar trebui să fim surprinşi. Am fost avertizaţi că aceasta este o problemă foarte DURĂ pe care neuroştiinţele încearcă să o rezolve. Modelul lui Edelman este plin de detalii, dar are puţine specificităţi. După cum a remarcat Daniel Dennett, Edelman repetă adesea idei elaborate de alţii fără a părea să aprecieze legăturile, iar proza sa poate fi de-a dreptul obscură. Un comentariu des repetat cu privire la model a venit de la cunoscutul biolog Gunther Stent,

Nu mă consider prea prost. Sunt profesor de biologie moleculară şi preşedinte al secţiei de neurobiologie a Academiei Naţionale de Ştiinţe, deci ar trebui să înţeleg. Dar nu înţeleg.

Într-adevăr.

Există un alt aspect al abordării lui Edelman care pare problematic. Atunci când a fost împins să fie mai specific cu privire la fundamentele neuronale ale conştiinţei, el a avut tendinţa de a îmbrăca discursul cu referiri constante la „amintiri comparative recurente" care, în modul său de gândire, erau legate de evenimente perceptive curente, în curs de desfăşurare. Dacă luăm acest lucru literal, aşa cum insista Edelman, s-ar părea că un nou-născut, lipsit de aceste amintiri comparative recurente, nu are nici conştiinţă. Pot aprecia o poziţie care susţine că experienţele fenomenale ale bebeluşului proaspăt născut nu sunt deosebit de bogate în conţinut mental şi că nu există prea multă conştiinţă de sine, dar există cu siguranţă experienţe, percepţii, conştiinţă. După cum s-a menţionat, nou-născuţii se pot deosebi pe sine de ceilalţi şi se ştie că se concentrează asupra feţelor, răspund la

46 Acest proces este similar cu modul în care se formează aşa-numitele „ansambluri de celule Hebbian". Donald Hebb, un neuropsiholog canadian, a introdus noţiunea de ansamblu celular în cartea sa fundamentală din 1949, *The Organization of Behavior.* În urmă cu aproape 70 de ani, Hebb a prezis că, odată cu repetarea datelor de intrare sau a acţiunii, grupurile de neuroni vor ajunge să formeze unităţi care se activează la unison sau, conform frazei devenite celebre (cunoscută sub numele de Legea lui Hebb), „neuronii care se activează împreună se conectează împreună".

vorbire, dau dovadă de atenție selectivă și manifestă simpatii și antipatii. Lucrări recente arată, în mod remarcabil, că fetușii din trimestrul al treilea răspund la modele punctuale care sunt de natură asemănătoare fețelor. Psihologul Vincent Reid și colaboratorii de la Universitatea Lancaster au folosit stimulente intense care puteau pătrunde în uter pentru a oferi fătului diferite modele vizuale. În cazul în care au fost prezentate într-un model care capta configurația unei fețe, sugarii au putut fi văzuți (pe înregistrările cu ultrasunete) urmărindu-l în timp ce se mișca. Același set de stimuli cu susul în jos a fost ignorat.

Edelman a recunoscut aceste stări interne ca făcând parte din grupul de valori înnăscute, dar pare să nu le considere caracteristici ale conștiinței - și acest lucru creează o problemă pentru modelul său. Dacă nou-născutul este lipsit de conștiință, atunci undeva în timpul dezvoltării și creșterii, sugarul/copilul/adultul trece de la o ființă lipsită de conștiință la una care o experimentează. Deoarece modelul lui Edelman nu reușește să se confrunte cu această dificultate, el, la fel ca multe alte teorii, este ridicat pe piedestalul dilemei emergentiștilor - care va primi o expunere mai dezvoltată mai jos.

Daniel Dennett.

Dennett, unul dintre cei mai influenți filosofi în viață, se luptă cu problema minții în 18 cărți și nenumărate articole, toate legate între ele de firul evoluției darwiniste prin selecție naturală. În cea mai recentă și mai lungă lucrare a sa, *From Bacteria to Bach and Back: The Evolution of Minds* (Evoluția minții), el susține că a corectat erorile din eforturile anterioare și a pus cap la cap lunga poveste de la primele forme de viață („bacteriile” din titlu), până la strălucirea și creativitatea speciei noastre („Bach”) și „înapoi” la un mod de a înțelege totul ca fiind fundamentat în termeni pur materialiști.

Abordarea lui Dennett nu este o teorie a apariției minții, ci mai degrabă un discurs extins despre tot ceea ce contează în știință și filosofie, la care poate apela pentru a picta o imagine coerentă a modului în care mințile umane au apărut începând de la început, supa prebiotică din care a apărut viața. În general, nu am nicio problemă cu prezentarea evenimentelor de către Dennett, cu numeroasele sale incursiuni într-o gamă largă de descoperiri științifice sau cu proza strălucită cu care își prezintă adesea opiniile. De exemplu, așa cum am menționat mai devreme, el face o distincție utilă între competență și înțelegere. Prima reprezintă, în opinia sa, noțiunea de moduri adaptative și eficiente de comportament care permit organismelor să facă față capriciilor și contingențelor lumii în care trăiesc. Acestea generează rutine care sunt executate fără conștientizare, fără nicio înțelegere fenomenală a ceea ce se întâmplă. Pe scurt, ele reflectă o competență care este lipsită de orice motive sau temeiuri pentru înțelegere - ceea ce necesită înțelegere.

Cred că această distincție este utilă, deși din motive oarecum diferite. Este un alt mod de a exprima o distincție pe care am făcut-o pentru prima dată în anii 1960, când am raportat o serie de studii care arătau că învățarea caracteristicilor structurale ale domeniilor complexe poate avea loc inconștient sau, folosind terminologia pe care am preferat-o, implicit. După cum au arătat cercetările ulterioare, *învățarea implicită* este foarte diferită de *învățarea explicită*, care are loc într-un cadru cognitiv de sus în jos, modulat în mod conștient. Dobândirea, stocarea și utilizarea implicită a cunoștințelor are loc rapid, este procedurală, automată și modulată de centre cerebrale care sunt vechi din punct de vedere evolutiv. Aceasta este partea „rapidă” în exegeza lui Daniel Kahneman din 2011 pe această temă, *Thinking, Fast and Slow.* Este, de asemenea, un proces care este robust în fața tulburărilor neurologice și/sau psihologice, apare într-o formă pe deplin funcțională la naștere și rămâne relativ consecvent pe tot parcursul vieții. Sistemul implicit este, de asemenea, relativ independent de funcționarea intelectuală generală și prezintă mai puține variații interindividuale decât operațiunile și procesele explicite. Pe scurt, acesta prezintă o competență asemănătoare cu cea a lui Dennett, fără înțelegerea obișnuită observată în alte funcții cognitive umane.

Dennett nu abordează o distincție conexă pe care o consider importantă din punct de vedere epistemic, și anume, cea dintre **a ști că** și a ști ce. Michael Polanyi s-a concentrat asupra acestei comparații în cartea sa din 1958, *Personal Knowledge*. El a observat că există situații în care știm că știm ceva fără să știm ce anume știm. În competența lui Dennett (sau în cunoașterea mea implicită) există o bună parte din cunoașterea a ceea ce în absența cunoașterii a ceea ce. Poate că nu știu ce reguli gramaticale au fost încălcate atunci când citesc o propoziție neconformă, dar simt imediat că acestea au fost încălcate. Nu știu ce principii bio-neuromusculare subiacente funcționează atunci când conduc o mașină, dar știu că acestea funcționează corect și că știu să conduc. Dennett pare să trateze competența ca având loc fără nicio conștientizare a acelei caracteristici. Dar aceasta există. În mai multe dintre experimentele noastre privind *învățarea implicită*, participanții care nu au putut să ne spună ce știau despre structura afișărilor de stimuli au produs date la nivel de încredere care au arătat că știau că știu ceva relevant (a se vedea Reber, 1989,1993).

Când sunt întrebat de către profani ce am studiat atunci când făceam toată această muncă empirică, mi se pare cel mai ușor să le spun că a fost intuiția umană, circumstanțele în care ai învățat ceva și știai că știi, dar nu puteai explica ce știai sau cum (sau chiar când) ai învățat. Exemplele mele obișnuite sunt gramatica limbii cu care am crescut vorbind și procesele de socializare care le-au impregnat cunoștințele despre cultura în care trăiesc. Acest lucru este de obicei răspuns cu o înclinare a capului în cunoștință de cauză. De asemenea, primesc zâmbete de recunoaștere dacă vorbesc cu cineva care joacă jocuri sofisticate precum pokerul, bridge-ul, go-ul sau șahul, în care aceste momente apar cu regularitate. Practic, toți jucătorii de poker de succes, de exemplu, au învățat să urmeze acest tip de sfaturi îndepărtate, șoptite.[47]

Există o vastă literatură tehnică privind învățarea implicită, memoria implicită și intuiția, aproape toată bazată pe ideea că există funcții cognitive care au loc în afara procesării standard, de sus în jos.[48] Aceste sisteme de detectare a modelelor de jos în sus produc cunoștințe utile și eficiente și, mai important, arată că deținătorul acestor cunoștințe știe acest lucru. Atunci când cineva ne spune că „a avut o presimțire” cu privire la o mutare pe care adversarul său de la șah urma să o facă, sau un jucător de poker spune că a acceptat ultimul pariu al adversarului său pentru că „a avut o vagă presimțire” că era o cacealma, sau un jucător de jocuri video anunță că „era destul de sigur” care va fi următoarea mutare a unei arme a adversarului său sau ... Ei bine, ați prins ideea. Toate aceste cazuri sunt situații în care există un „știu că” în absența unui „știu ce”. Chiar și speciile unicelulare prezintă dovezi pentru acest tip de procesare. Revenim la descrierea lui Jennings despre *Stentor roeselii*, pe care am discutat-o mai devreme. După mai multe încercări cu particule toxice aruncate în pâlnia sa de admisie, acesta știa, avea o reprezentare internă, că sunt neplăcute și trebuie evitate și a acționat pe baza acestei cunoștințe pentru a se proteja. Dar, bineînțeles, corpul său celular singular, fără neuroni, nu avea nimic asemănător unei reprezentări mentale a ceea ce știa.

Dar ceea ce ne deosebește pe Dennett și pe mine este faptul că el caută punctul (punctele?) din dezvoltarea evolutivă în care această competență, pe care o atribuie tuturor organismelor, începând cu cele mai simple (din nou bacteriile), face loc tipului de înțelegere a acestor lucruri la care mințile umane sunt atât de bune. Dennett nu dorește să înzestreze o specie unicelulară precum *Stentor* cu ceva asemănător unei vieți mentale reale. Eu vreau. Nu că aș crede că *Stentor* gândește în vreun mod epistemic interesant. Dar cred că tipurile de gândire în care ne angajăm, în care se angajează corvidele, cetaceele, cefalopodele, bondarii, caninii, primatele neumane și multitudinea de alte specii ale căror talente cognitive vor fi cu siguranță descoperite în anii următori, se va dovedi că toate își au rădăcinile în acele versiuni primordiale. Abordarea lui Dennett se încadrează în cadrul biofilosofic actual, dominant, în care funcțiile cognitive umane (și poate și cele ale altor specii) sunt

47 Sunt jucător de poker încă din timpul facultății și, de asemenea, am scris trei cărți și câteva sute de articole despre acest joc și despre jocurile de noroc în general. Dacă cineva este curios, cartea mea *Poker, Life, and Other Confusing Things*, o colecție de eseuri despre, evident, poker și viață, este un loc bun de început. În altă parte, am prezentat un nou cadru de abordare a aspectelor juridice, legislative și etice legate de jocurile de noroc (Reber, 2012b). Uneori trebuie doar să te îndepărtezi de știința conștiinței și să te implici în alte domenii.

48 A se vedea Reber (1993), Stadler și Frensch (1998) și French și Cleeremans (2002) pentru detalii privind cercetările teoretice și empirice anterioare. A se vedea Tin (2014) pentru o prezentare generală a descoperirilor și extinderilor mai recente ale modelului în alte domenii ale psihologiei cognitive și sociale.

tratate ca și cum ar fi de un tip separat de cele întâlnite la celelalte specii care populează planeta. Abordarea **Bazei Celulare a Conștiinței** le înțelege pe toate ca fiind situate pe un continuum și le tratează pe toate ca pe simboluri distincte ale unui singur tip.

În cele din urmă, există o altă dificultate pe care o am cu abordarea lui Dennett, iar poziția mea este împărtășită de mulți alții. Este vorba de refuzul său hotărât și absolut de a recunoaște că conștiința are ceva care se apropie de un statut ontologic acceptabil, că are vreun rol cauzal de jucat în comportament sau gândire. Pentru el, aceasta este o „iluzie a utilizatorului" pe care o creăm pe măsură ce observăm, recunoaștem sau devenim conștienți de numeroasele lucruri minunate, pătrunzătoare, creative și delicate pe care le facem, le spunem, le gândim și le simțim. El susține că am fost păcăliți să credem în propria noastră perspectivă la persoana întâi. Simțim căldura de la sobă și ne îndepărtăm de ea. Ascultăm cel de-al doilea concert brandenburgic al lui Bach și simțim un fior de frumusețe și admirație. Vedem un banner nazist și simțim repulsie. Dar, susține el, nu știm cu adevărat ce se întâmplă aici, nici în sistemele noastre perceptive, nici în mușchii noștri, nici în creierul nostru. Noi doar reacționăm. Credem că avem parte de o experiență fenomenală autentică, care este o manifestare sigură a conștiinței, dar nu este așa. Este doar o iluzie a utilizatorului. De fapt, noi doar ne comportăm.

Dennett pare să nu fi fost în stare să renunțe la Ryle.[49] O mare parte din această temă conform căreia conștiința nu este decât o iluzie a utilizatorului derivă din pozițiile antidualiste ale lui Gilbert Ryle expuse în volumul său din 1949, *The Concept of Mind*. Ryle s-a străduit să mențină behaviorismul care a dominat gândirea și teoria în acele decenii, ridicând în același timp puțin pânzele cortului și încorporând mintea și funcțiile mentale - atât timp cât acestea și-au păstrat cârligele conceptuale ferm atașate de modelele materialiste. Ryle, la fel ca Dennett, se simțea confortabil cu termenii cognitivi, precum și cu procesele și funcțiile mentale. Spre deosebire de contemporanii săi, precum Clark Hull, Edwin Ray Guthrie și B. F. Skinner, care au îmbrățișat behaviorismul radical, Ryle a sărbătorit gândirea controlată de sus în jos, așa cum face Dennett. Dar, din motive pe care, sincer, le consider misterioase, în cadrul său conceptual nu pare să existe loc pentru tratarea *qualia* ca fenomenal real sau pentru o conștiință ontologică.

În mod interesant, psihologii britanici David Oakley și Peter Halligan au prezentat recent un argument similar, care, în esență, califică conștiința drept un epifenomen fără puteri cauzale. În cuvintele lor,

În ciuda experienței subiective convingătoare a autocontrolului executiv, susținem că „conștiința" nu conține procese de control de sus în jos și că „conștiința" nu implică nicio relație executivă, cauzală sau de control cu niciunul dintre procesele psihologice familiare atribuite convențional acesteia. În opinia noastră, procesarea psihologică și produsele psihologice nu se află sub controlul conștiinței - conștiința personală este analogă curcubeului care însoțește procesele fizice din atmosferă, dar nu exercită nicio influență asupra acestora.

Dacă aș putea simți că există un câștig epistemic din separarea conștiinței de restul sensibilității, aș fi mai înțelegător față de eforturile lui Dennett și ale lui Oakley și Halligan. Dar nu văd niciunul. Nu înțeleg cum înțelegerea noastră a condiției umane, a minții umane (Dennett este de acord cu această noțiune de minte, dar nu și cu conștientizarea ei de către noi) sau cum modelele sau teoriile noastre despre locul ei în arborele vieții sunt îmbunătățite prin eliminarea conștiinței. De fapt, a urma linia filozofică a lui Dennett (și a lui Oakley și Halligan) pare o tortură. Trebuie să treci peste natural pentru a te angaja în nenatural, să defenestrezi bunul simț, să abandonezi limbajul ușor al discuțiilor cotidiene despre minte. Apreciez că toți încearcă să evite genul de enigme filosofice care i-au împins pe alții, precum Chalmers, Nagel sau Koch și Tononi, să adopte poziții dualiste sau panpsihice, dar, după cum susțin aici, există o altă cale către materialism care păstrează conștiința ca concept viabil. Pentru a ajunge acolo, Dennett (ca și Oakley și Halligan) ar trebui să facă doar o mică ajustare în gândirea sa și să accepte că competența sa are o caracteristică de conștientizare de

49 Dennett și-a luat doctoratul în 1965 la Oxford sub îndrumarea lui Gilbert Ryle.

acest tip şi că se află pe acelaşi continuum ontologic ca şi înţelegerea sa - şi că ambele încep cu acele bacterii nenorocite de care recunoaşte că este intrigat.

Antonio Damasio.

Atacul general al lui Damasio asupra problemelor legate de conştiinţă şi minte, argumentat în cartea sa revoluţionară din 1994, *Descartes' Error* (Eroarea lui Descartes), s-a îndepărtat de majoritatea celorlalţi prin accentuarea emoţiilor, sentimentelor şi a unei componente corporale. Aceasta a fost o schimbare importantă în centrul atenţiei, deoarece eforturile anterioare, după cum am observat, tindeau să pună accentul pe mecanismele şi procesele cognitive şi perceptive, neglijând, la fel ca Descartes, afectivitatea, emoţiile şi motivaţia.

În lucrări ulterioare, cum ar fi *The Feeling of What Happens* şi *Self Comes to Mind*, Damasio şi-a dezvoltat modelul mai detaliat, concentrându-se pe relaţiile dintre elementele cognitive şi cele afective, identificând structuri şi căi cerebrale specifice care joacă un rol în aceste procese şi, ceea ce este important, legându-le de funcţiile şi reprezentările corporale. În modul de gândire al lui Damasio, trebuie, de asemenea, să se facă distincţia între emoţii şi sentimente. Primele sunt, în concepţia sa lexicografică, reacţii automate, neaşteptate la circumstanţe. Sentimentele, pe de altă parte, necesită o interpretare, o evaluare a stării emoţionale.[50] Într-un interviu acordat lui Jason Pontin în *MIT Technology Review* (17 iunie 2014), Damasio a explicat astfel:

Trebuie să separăm componenta care provine din acţiuni de componenta care provine din perspectiva noastră asupra acelor acţiuni, care este sentimentul. În mod curios, acesta este şi locul în care apare sinele şi conştiinţa însăşi. Mintea începe la nivelul sentimentelor. Atunci când ai un sentiment (chiar dacă eşti o creatură foarte mică), începi să ai o minte şi un sine.

Toate aceste elemente ale conştiinţei sunt, totuşi, înfăşurate în conştiinţa de sine. Pentru Damasio, simpla experienţă nu este suficientă. Ceea ce contează este o conştiinţă de sine a ceea ce înseamnă să ai o experienţă. După cum spune el,

în absenţa conştiinţei, viziunea personală este suspendată; nu ştim despre existenţa noastră şi nu ştim că orice altceva există.

Conştiinţa lui Damasio este un tip diferit de întreprindere faţă de cele propuse de Dennett, Edelman sau Hameroff şi Penrose - deşi a remarcat că nu este incomod cu modul în care Edelman abordează subiectul. Damasio nu susţine o ipoteză specifică cu privire la originile minţii, ci prezintă o exegeză extinsă asupra funcţiei mentale umane şi a componentelor neuronale, corporale şi emoţionale care interacţionează în moduri extrem de complexe pentru a o produce. Damasio este mai degrabă un neuroştiinţific care comentează despre minte decât un apărător al unei viziuni specifice asupra modului exact în care creierul uman creează mintea umană. Damasio este, de asemenea, dispus să ia în considerare ideea că rădăcinile evolutive ale conştiinţei ar putea fi văzute în funcţiile microbiologice ale speciilor unicelulare - aşa cum am menţionat mai devreme, el a fost unul dintre semnatarii scrisorii adresate lui Koch de către Sacks et al. În schimburile personale pe care le-am avut cu el, am înţeles că încă mai suspectează că ceva asemănător ţesutului neural, un sistem nervos oricât de primitiv, este necesar pentru adevărata conştiinţă. În acest caz, poziţia lui Damasio, ca şi a altora pe care le-am abordat, va avea probleme cu *emergentismul*.

---

50 Există o paralelă interesantă între distincţia dintre emoţii şi sentimente pe care Damasio o face şi cea pe care Dennett o face între competenţă şi înţelegere. Ambele presupun un grup de funcţii primitive care operează în mare măsură independent de conştiinţă şi o altă categorie, mai sofisticată, de acte mentale marcate de conştientizare, interpretare şi înţelegere. Din câte ştiu eu, niciunul dintre ei nu a recunoscut aici perspectiva celuilalt, poate pentru că Damasio vede conştiinţa ca pe un element central al minţii, în timp ce Dennett o tratează ca pe un resturi epifenomenale ale operaţiunilor mentale inconştiente.

Giulio Tononi şi Christoph Koch.

Tononi şi Koch sunt amândoi specialişti în neuroştiinţe cu interese vaste în subiecte care variază de la vedere până la somn şi, desigur, la conştiinţă. Colaborarea lor a început la scurt timp după ce, în 2004, Tononi a prezentat *Teoria integrată a informaţiei* (IIT) ca model formal al minţii şi conştiinţei. IIT se bazează pe teoria fundamentală a informaţiei a lui Claude Shannon, veche de 70 de ani (a se vedea Shannon, 1948), care presupunea că, ori de câte ori incertitudinea din orice sistem este redusă (de exemplu, prin efectuarea unei alegeri, prin aruncarea unei monede, prin apăsarea unui buton), se obţin informaţii, se dobândesc cunoştinţe.

Tononi a extins domeniul factorilor şi a introdus o măsură formală, teoretică, desemnată prin litera greacă **phi** (Φ), a oricărui sistem bazat pe gradul în care fiecare dintre elementele sau componentele separate ale sistemului sunt legate sau integrate în alt mod în relaţii funcţionale între ele. Adică, <Φ> exprimă nu doar cantitatea de informaţie produsă de orice acţiune a fiecăreia dintre componentele independente ale oricărui sistem arbitrar, ci include şi toate interacţiunile posibile dintre ele. De fapt, dacă un sistem are doar părţi, dar nu are interacţiuni între ele, <Φ> a sistemului = 0. Cu cât sunt mai mulţi factori şi cu atât mai extins sunt interconectaţi funcţional, cu atât valoarea lui <Φ> este mai mare şi — aici începe distracţia — în versiunea Tononi-Koch a IIT, <Φ> se presupune a fi o măsură a conştiinţei.

Majoritatea analizelor lor privind <Φ> şi modul de măsurare a acesteia se concentrează asupra sistemului nervos uman. Ei observă că cerebelul, în ciuda faptului că are mult mai mulţi neuroni individuali decât cortexul cerebral, joacă un rol mai mic în subiectivitate şi conştiinţă. Ei interpretează acest fapt ca fiind în concordanţă cu IIT pe motiv că neuronii cerebeloşi sunt mult mai puţin integraţi şi interconectaţi decât cei corticali. Pierderea conştienţei în timpul unei convulsii este considerată ca fiind rezultatul creşterii sincroniei şi al scăderii concomitente a integrării diferitelor componente corticale din creier care însoţesc astfel de atacuri. Argumente similare sunt aduse şi în cazul altor stări în care conştiinţa este diminuată sau lipseşte cu desăvârşire, cum ar fi sub diferite niveluri de anestezie sau în diferite stadii ale somnului. Într-o lucrare recentă de prezentare generală (Koch. Massimini, Bolv şi Tononi, 2016), aceştia concluzionează că, deşi nicio zonă din creier nu este esenţială pentru conştiinţă, există corelaţii neuronale care joacă un rol semnificativ şi acestea nu se află în căile frontoparietale, aşa cum mulţi au susţinut, ci se află în principal în cortexul senzorial posterior.

Dacă modelul Tononi-Koch s-ar încheia aici, cu un cadru care să ghideze explorarea corelaţiilor neuronale dintre creier şi minte şi cu o modalitate de principiu de măsurare a conştienţei, am avea o abordare interesantă, una care împărtăşeşte caracteristici cu cea a lui Damasio şi Edelman, adăugând în acelaşi timp proprietăţile formale ale <Φ> care, deşi confruntată cu probleme pe care le vom aborda mai târziu, are potenţial ca modalitate de a pune o metrică cantitativă pe capacitatea mentală. Cu toate acestea, Tononi şi Koch nu numai că se aventurează dincolo de descoperirea corelaţiilor neuronale-comportamentale şi de măsurarea nivelurilor de conştiinţă, ci merg pe două căi pe care le-am identificat deja ca fiind, în cel mai bun caz, neproductive şi, mai probabil, confuze din punct de vedere epistemic: funcţionalismul independent de hardware şi o versiune modificată a panpsihismului.

Aceste probleme sunt create prin tratarea lui <Φ> ca o măsură a conştienţei şi prin presupunerea că orice sistem cu elemente care interacţionează are un Φ > 0. Odată făcută această mişcare, zarurile sunt aruncate, runele sunt sculptate, mâna lor este forţată. Abordarea IIT devine astfel o altă variantă a panpsihismului în care chiar şi sistemele moderat de complexe, cum ar fi computerele sau maşinile care se conduc singure, sau chiar fotodiodele şi termostatele, sunt tratate ca având un anumit grad de conştiinţă. Informaticianul Scott Aaronson[51] a criticat această poziţie. IIT este, a scris el, o teorie care,

51 Aaronson are un blog popular pe care îl numeşte Shtetl-Optimized, unde poate fi găsită o serie de postări critice la adresa IIT (http://www.scottaaronson.com/ blog/?s=Tononi).

prezice în mod inevitabil cantităţi uriaşe de conştiinţă în sisteme fizice pe care nicio persoană sănătoasă nu le-ar considera deloc „conştiente”*:*
într-adevăr, sisteme care nu fac altceva decât să aplice un cod de paritate cu verificare de densitate redusă sau alte transformări simple ale datelor lor de intrare. (Evidențierea textului cu caractere aldine în original.)

IIT impune constrângerea că părțile din sistem trebuie să fie integrative, cu bucle de feedback și caracteristici de alegere funcțională. Într-un efort de a evita tipurile de panpsihism fără legătură pe care le-am observat la alții, precum Max Velmans și Philip Goff, Koch și Tononi susțin că IIT nu se aplică la setările care nu interacționează, cum ar fi o grămadă de nisip sau chiar instrumente complexe precum camerele foto. Dar acest lucru ridică o serie de întrebări. Dacă o grămadă de nisip este scuturată și boabele, având diametre ușor diferite, suferă o sortare verticală, ceea ce presupune interacțiunea lor, grămada de nisip a devenit brusc conștientă? Dacă o cameră este integrată într-un vehicul în mișcare cu senzori complecși, sisteme de feedback și componente capabile să ia decizii, cum ar fi o mașină care se conduce singură, a dezvoltat aceasta sensibilitate? Dacă da, niciunul nu ne-a dat vreun indiciu în acest sens.

John Searle, așa cum ne așteptam, și-a îndreptat, de asemenea, privirea filozofică asupra modelului Tononi-Koch, observând că utilizarea teoriei informației ca fundament al IIT este o alegere ciudată, deoarece însuși conceptul de informație are sens doar atunci când există deja o minte conștientă care să verifice că există, într-adevăr, informații în sistem. Sensibilitatea despre care Koch și Tononi susțin că este exprimată de orice lucru care posedă informații integrate, cum ar fi termostatele, întrerupătoarele de lumină sau computerele, este determinată ca fiind fenomenală doar pentru că cineva cu experiențe fenomenale a interpretat-o astfel.

Această confuzie amintește de vechea enigmă care apare în jurul celei de-a treia clase (sau, pentru unii, în clasa a 1-a): „*Dacă un copac cade în pădure, dar nu este nimeni prin preajmă care să-l audă, face oare un zgomot?*” Motivul pentru care această întrebare îi nedumerește pe mulți este că nu precizează despre ce fel de sunet este vorba. Interjecția „dar nu este nimeni în jur” ne face să ne gândim la un sunet perceput, la un sunet interpretat, la un sunet de psiholog mai degrabă decât la un sunet de fizician.

Sunetul are proprietăți fizice. Acesta este compus din condensări și rarefacții ale moleculelor dintr-un mediu care se propagă de la o sursă în unde. Atunci când aceste unde sonore se află într-o anumită gamă de frecvențe și ajung deasupra unui prag de intensitate specificabil pe o ureche care este conectată la un creier cu structuri neuronale adecvate, are loc o experiență mentală, subiectivă a sunetului. Primul sens este un analog al sensului original al informației al lui Shannon. A doua este un analog al transformării Tononi-Koch a acesteia în IIT. Nu poți avea sunet fenomenal fără o entitate care să îl audă și să îl interpreteze. Nu puteți obține informații instanțiate conștient fără un agent care să dea sens parametrilor fizici. Dar aceste condensări și rarefacții ale moleculelor într-un mediu există și nu contează deloc cine, care sau ce locuitor al pădurii se află sau nu prin preajmă.

Legătura dintre informație, măsurată ca Φ, este doar atât, o măsură a informației, iar teoriile bazate pe informație sunt în mod inerent circulare; ele sunt toate legate de observator. Așa cum a spus scriitorul de știință John Horgan,

Conceptul de informație nu are sens în absența a ceva care să fie informat - adică un observator conștient capabil de alegere.

Modelul Tononi-Koch al rădăcinilor conștiinței pur și simplu nu poate „trece” fără a avea deja un agent conștient, astfel încât, în mod logic, nu ne poate oferi o explicație a originilor conștiinței. Apoi, există problema practică a calculării <Φ>, care implică contabilizarea nu numai a fiecărui element individual, ci și a fiecărei interconexiuni dintre acestea. Tononi și Koch observă că numărul de moduri în care creierul unui nematod (o specie de vierme cu doar 302 neuroni) poate fi analizat este numărul astronomic 10 urmat de 467 de zerouri. Pe scurt, intractabil din punct de

vedere computațional. Creierul uman are undeva între 85 și 100 de miliarde de neuroni și peste 100 de trilioane de sinapse. Obținerea unei valori de <Φ> pentru *Homo sapiens* este, în toate scopurile practice, o imposibilitate de calcul. După cum a remarcat psihologul cognitiv Gary Marcus, acest lucru face din <Φ> puțin mai mult decât o curiozitate intelectuală, nu o măsură utilă a conștiinței - sau, de altfel, orice altceva.

Din perspectiva psihologiei populare pe care am adoptat-o anterior, este cu siguranță rezonabil să ne întrebăm de ce, dacă Tononi și Koch au dreptate, laptopurile sau telefoanele mobile nu au prezentat conștiință, conștientizare. Până în prezent, nu există niciun indiciu de subiectivitate, de experiență fenomenală din partea niciunuia dintre ele. Este foarte probabil ca <Φ> pentru un dispozitiv atât de complex precum un computer de birou să fie mai mare decât cel al unui *Paramecium*. Primul nu prezintă niciun semn de animism, nicio agenție; cel de-al doilea destul de mult din ambele. Ce spuneți de supercomputerele care sunt utilizate pentru a analiza datele meteorologice, pentru a modela schimbările climatice, pentru a efectua calcule pentru experimente în astrofizică și mecanică cuantică? Sau internetul, cu raza sa globală și miliardele sale de noduri? La fel ca sistemul de telefonie din povestea științifico-fantastică a lui Arthur C. Clark pe care am menționat-o în capitolul 1, acesta ar trebui să aibă un număr suficient de mare pentru conștiința de sine. În cartea sa *Consciousness: Confessions of a Romantic Reductionist* (Confesiunile unui reducționist romantic), Koch își pune această întrebare, dar nu pare îngrijorat de lipsa dovezilor în timp ce meditează la posibilitatea ca, într-o bună zi, internetul să ni se poată prezenta ca fiind conștient.

În cele din urmă, deși nu sunt chiar captivat de IIT sau de direcția în care Tononi și Koch l-au dus, am numai cuvinte de laudă pentru Tononi și Koch ca cercetători și scriitori, neuroștiințifici cognitivi productivi și respectați. Laboratorul lui Tononi de la Universitatea din Wisconsin este un centru de vârf pentru studiul somnului, al funcțiilor sale, al fundamentelor sale genetice, al relației sale cu conștiința și al corelațiilor neurologice ale tulburărilor de somn. Koch este președinte și director științific al Institutului Allen pentru Știința Creierului din Seattle, conducând în același timp laboratorul său de la Institutul de Tehnologie din California. Timp de peste 30 de ani, acesta a adus o contribuție inovatoare constantă în domeniul neuroștiințelor, de la modelarea sistemului vizual, la simulări ale acțiunilor țesutului neural, la explorarea corelațiilor cerebrale ale conștiinței. În mod interesant, atât Tononi, cât și Koch au colaborat cu laureați ai Premiului Nobel: Tononi cu Gerald Edelman, Koch cu Francis Crick.

Problema nu este cu munca lor empirică, ci cu eforturile lor de a rezolva problema dificilă. Se pare că fiecare om de știință bun care urmărește acest *graal* fals ars, începând cu conștiința umană, ajunge într-un loc incoerent, gol ontologic. Un avantaj suplimentar al poziției **CBC** este că ocolește aceste puncte slabe, cele care au condus unii cercetători foarte serioși și respectați într-un misterianism nedorit. Mai multe despre această problemă mai târziu.

Pe lângă aceste eforturi, mai mulți cercetători și filosofi au intrat pe câmpul de luptă, în unele cazuri cu puncte de vedere destul de noi și interesante. Merită să aruncăm o privire rapidă și critică asupra câtorva dintre aceștia.

Julian Keenan (împreună cu Gordon Gallup și Dean Falk).

În cartea lor din 2003, *The Face in the Mirror*, ei, la fel ca Damasio, susțin că conștiința de sine, autoconștiința, caracteristicile autoreferențiale ale sensibilității sunt cheia înțelegerii minții umane. Ei remarcă faptul că puține specii pot trece de sarcina „fața în oglindă”, pe care o consideră o condiție pentru ca o specie să fie considerată cu adevărat posesoare a unei conștiințe personale. Această sarcină, dezvoltată de Gallup, este una în care un semn, precum un punct roșu, este plasat pe animal într-un loc în care nu poate fi văzut decât într-o oglindă. Dacă animalul atinge punctul din reflexie, conform interpretării standard, nu a reușit să se recunoască pe sine. În cazul în care

animalul se atinge acolo unde se află punctul (de exemplu, pe frunte), acest lucru indică faptul că animalul s-a recunoscut în oglindă şi, astfel, manifestă conştiinţă de sine.

Într-o serie extinsă de studii în mai multe laboratoare diferite, doar câteva specii au trecut testul, printre care unele maimuţe mari, inclusiv cimpanzeii şi bonobo, cetaceele şi, posibil, elefanţii. Unele dintre cele mai interesante studii au fost realizate de expertul în cetacee Diana Reiss de la Colegiul Hunter al Universităţii oraşului New York. Într-un experiment, delfinii sunt marcaţi în timp ce interacţionează cu un om. Imediat după aplicarea mărcii, aceştia înoată până la o oglindă care a fost plasată în apă şi îşi întorc corpul pentru a putea vedea ce s-a întâmplat. Este clar că nu numai că trec testul, dar au şi un interes activ pentru marcaj şi ştiu exact unde să se uite.

Deşi este o abordare interesantă şi diferită a conştiinţei, teoria suferă de două probleme. În primul rând, Keenan şi colaboratorii săi presupun că recunoaşterea sinelui în sarcina oglinzii este un marker fiabil al conştiinţei de sine. Pur şi simplu nu este clar că această deducţie este susţinută. A fi conştient de sine implică mai mult decât recunoaşterea propriului sine fizic; implică o capacitate reflexivă de a-şi cunoaşte sinele esenţial sau caracterul, de a fi conştient de propriile sentimente şi emoţii. Care nivel de conştientizare este esenţial? În ceea ce priveşte celelalte specii care au trecut testul, suntem din nou în vechea noastră „capcană Nagel". Avem doar o perspectivă de a treia persoană asupra animalului şi nu avem o înţelegere reală a experienţelor sale subiective. S-ar putea foarte bine ca autorecunoaşterea să fie necesară, dar nu suficientă pentru tipurile de conştiinţă pe care le au oamenii, dar nu văd nicio modalitate de a ne da seama.

În al doilea rând, prin limitarea conştiinţei autentice doar la acele specii care sunt conştiente de sine (indiferent de modul în care o evaluăm), ei se expun unei cerinţe emergente severe. De ce, cum şi când, în termeni evolutivi, a apărut acest aspect conştient de sine al vieţii mentale? Şi ce caracteristici şi proprietăţi are acesta care îl fac să fie determinantul singular al conştiinţei? Cred că acestea nu numai că sunt probleme insurmontabile, dar nici măcar nu sunt cele care trebuiau să ne fie prezentate. Este arbitrar să presupunem că conştiinţa ontologic valabilă se află doar în domeniul privilegiat al conştiinţei de sine. De ce ar trebui să fie aşa? Aceasta nu este atât o propunere ştiinţifică, cât una lexicografică. Ea oferă doar o definiţie a conştiinţei - şi încă una destul de restrictivă. Dar un element al abordării lor este intrigant. Keenan et al. iau în considerare posibilitatea ca zonele neuronale critice să se afle în emisfera dreaptă, adesea neglijată. La fel ca Damasio şi alţii, ei citează sindroame clinice pentru a le susţine.

Julian Jaynes.

Fără îndoială, cea mai neobişnuită teorie este cea pe care Jaynes a introdus-o în 1976 în cartea sa, *The Origin of Consciousness in the Breakdown of the Bicameral Mind.* În modul de gândire al lui Jaynes, conştiinţa a fost mai mult sau mai puţin inventată în urmă cu aproximativ 3 000 de ani, ca strategie mentală, atunci când mintea „bicameralăˮ - care consta din cele două emisfere ale creierului conectate din punct de vedere neurologic, dar independente din punct de vedere conceptual - a început să îşi piardă capacitatea de a funcţiona eficient pe măsură ce culturile au devenit mai complexe şi mai greu de navigat.

Înainte de a renunţa la bicameralismul cognitiv, despre care Jaynes susţinea că produce o minte nonconştientă incapabilă de funcţii metacognitive şi o lipsă de reprezentări autobiografice, oamenii funcţionau într-un mod pe care el îl asemăna cu schizofrenia, în care o emisferă „auzea voci" venind din cealaltă. Jaynes a obţinut sprijin pentru această teorie ciudată prin aruncarea unei plase foarte largi, inclusiv printr-o analiză neortodoxă a *Iliadei* lui Homer, căutând dovezi care să-i susţină ideea că minţile bicamerale erau incapabile de introspecţie. Puţini sunt cei care iau în serios modelul său, iar acesta a găsit puţin sprijin în neuroştiinţe, însă îndrăzneala lui Jaynes şi amploarea sa ştiinţifică au atras un interes considerabil. Cartea a devenit un bestseller, a fost recenzată pe scară largă şi chiar nominalizată pentru un *National Book Award*. Influenţa sa a continuat în rândul publicului larg şi chiar a crescut în timp. La aproximativ 20 de ani de la publicarea originală, a fost înfiinţată Societatea Julian Jaynes, care numără astăzi aproximativ 400 de membri activi. Fascinaţia

generală pentru această carte, singura publicată de Jaynes, a fost exprimată cel mai bine de biologul evoluționist Richard Dawkins:

Este una dintre acele cărți care este fie un gunoi total, fie o operă de geniu desăvârșit, nimic la mijloc! Probabil prima variantă, dar îmi mențin pariul.

Maxine Sheets-Johnstone.

Sheets-Johnstone a ajuns la subiectul conștiinței din lumea dansului și, deloc surprinzător, adoptă ceea ce este cunoscut sub numele de „întoarcerea corporală". În lucrarea *The Primacy of Movement* (1999/2011), ea a susținut că poziția dominantă în filosofia minții, care a pus accentul pe funcțiile cognitive și mentale, neglijând întruparea, și-a creat probleme și a generat enigme precum *Hard Problem* (Problema dificilă). Pentru Sheets-Johnstone, conștiința a evoluat ca o reprezentare internă a unui corp, un organism dinamic, cinetic, locomotor, care navighează într-o lume complexă, în continuă schimbare. Pentru ea, *qualia* nu există ca experiențe senzoriale sau mentale punctuale, ci ca reprezentări integrate care contribuie la capacitatea de a supraviețui și de a se deplasa cu succes. Perspectiva lui Sheets-Johnstone diferă de majoritatea celor pe care le analizăm în această secțiune prin faptul că ea concluzionează fără ambiguitate că această conștiință bazată pe mișcare este la fel de veche ca viața însăși. În lucrarea sa cea mare, ea identifică, în mai multe locuri, bacteriile ca organisme pe deplin conștiente.

Așa cum Damasio a susținut că teoriile conștiinței sunt condamnate să fie incomplete dacă neglijează rolul emoțiilor și al sentimentelor, Sheets-Johnstone susține că o soartă similară trebuie să aibă cele care nu reușesc să pună mentalul în fizic și să cartografieze modalitățile dinamice și complexe de mișcare. Cognitiviștii și filosofii minții, a susținut ea,

au fost atât de hipnotizați de creiere ... [încât] lumea mai largă a creaturilor din care fac parte oamenii este uitată, neglijată în mod grav sau distorsionată cu aroganță.

Multe (cele mai multe?) dintre aceste eforturi nu sunt atât încercări de a se adresa direct *problemei dificile*, cât de a caracteriza modul în care conștiința este exprimată la oameni și de a identifica rolurile pe care centrele, căile și/sau rețelele cerebrale le joacă în sensibilitate. Din nou, nu este nimic în neregulă cu acest mod de abordare a problemelor; este doar o abordare pentru a face știință bună, neuroștiință cognitivă. De asemenea, aceasta a produs unele perspective fascinante asupra relațiilor dintre structurile neuronale și comportamentele cognitive și afective. Cartea lui Stanislaus Dehaene din 2014, *Consciousness and the Brain: Deciphering how Brain Codes our Thoughts* (Conștiința și creierul: Descifrarea modului în care creierul ne codifică gândurile), este o prezentare admirabilă a acestor realizări. Aplaud eforturile și aștept cu nerăbdare descoperirile viitoare - deși nu anticipez că vreuna dintre ele va rezolva *problema dificilă*.

Căutarea filogenetică a primei apariții a conștiinței

În plus față de aceste eforturi, a existat un program paralel de cercetare, de natură mai mult istorică, care încearcă să identifice punctul din evoluție în care a apărut o specie[52] despre care se

52 Sau phylum, sau clasă, sau gen - adesea nu este deloc clar la ce nivel al taxonomiei filogenetice se concentrează un autor.

poate concluziona cu oarecare încredere că a posedat ceva asemănător unei forme primitive de conștiință sau că avea prezente caracteristici asemănătoare minții. Într-o lucrare din 2016, Stevan Hamad a formulat problema succint și, poate fără să vrea, a anticipat ceea ce se va dovedi a fi o problemă insurmontabilă.

[E]vident că organismele au minte, dintr-un motiv sau altul - sau cel puțin unele specii au minte: Care dintre ele? În mod clar, propria noastră specie are. Care altele? Și ce înseamnă să „ai o minte”?

Această abordare, așa cum trebuie să fie, este interdependentă de abordările descrise anterior, în care se caută corelații neuronale. Odată ce anumite centre și/sau căi neuronale sunt identificate în creierul uman ca fiind implicate în conștiința sau sensibilitatea umană, descoperirea servește ca o invitație la căutarea rădăcinilor lor evolutive. Odată ce s-a demonstrat în mod convingător că o specie prezintă o anumită formă de sensibilitate, descoperirea face practic obligatorie căutarea de structuri și forme similare sau analoge la om. Și, bineînțeles, atunci când se ajunge la concluzia că aceste căi sau structuri neuronale necesare nu sunt prezente la speciile țintă, se ajunge la concluzia că și aceste specii nu au acea formă particulară de sensibilitate. Am văzut cum s-a întâmplat acest lucru în cazul concluziei lui Brian Key conform căreia peștii nu simt durere fenomenală adevărată, deoarece nu au structurile neuronale corespunzătoare care mediază durerea la oameni și la alte specii.

Privind înapoi în timp în acest fel, făcând ceea ce pare a fi o ramură a ecologiei cognitive, poate avea unele implicații interesante și ridica ipoteze intrigante. De exemplu, după cum am menționat, în teoria lui Antonio Damasio, nucleul de bază al conștiinței este înrădăcinat în afecte, dar el face distincție între *emoții* și *sentimente*. Primele, susține el, sunt mai vechi din punct de vedere evolutiv și sunt prezente ca parte a experienței majorității formelor de viață, poate chiar și a speciilor unicelulare. Dar *sentimentele*, în taxonomia sa, necesită analiză, interpretare. Ele apar doar la speciile cu mecanisme neuronale mai sofisticate care permit analiza și interpretarea emoției.

Prin urmare, pentru cercetătorii care adoptă o poziție asemănătoare cu cea a lui Damasio, se vor căuta două noduri critice din punct de vedere evolutiv, unul care dezvăluie condițiile de apariție a experiențelor senzoriale și afective brute (emoții) și altul care permite o analiză mai de sus în jos a acestor experiențe și le permite să fie exprimate ca stări afective interpretate (sentimente). Primul punct va fi probabil mai devreme în timpul geologic decât cel de-al doilea, iar speciile care reflectă fiecare vor avea probabil arhitecturi neuronale diferite.

Din acest program de cercetare au rezultat o serie de informații fascinante despre funcțiile corticale și subcorticale, dar, desigur, nimic nu s-a apropiat de răspunsul la întrebările care se ascund în spatele *problemei dificile*. Este necesară o scurtă trecere în revistă a unora dintre cele mai interesante abordări.

<u>Todd Feinberg și Jon Mallatt</u>.

Într-o serie de cărți și lucrări recente, Feinberg, neurolog la Mount Sinai din New York, și antropologul Mallatt, în prezent la Universitatea din Idaho, au dezvoltat un cadru neurobiologic pentru ceea ce ei consideră a fi calitățile neuronale criteriale ale conștiinței. Ei au identificat patru trăsături care caracterizează ceea ce ei numesc conștiință neuro-ontologică:

- *<u>Referință</u>: Experiențele conștiente sunt despre (se referă la) lumea exterioară, corp sau stări afective, dar nu se referă la neuronii care produc experiențele.*
- *<u>Unitate mentală</u>: Conștiința este legată într-un câmp relativ unificat de conștiință, în contrast cu setul divizibil de neuroni individuali care o creează.*
- *<u>Cauzalitatea mentală</u>: Mintea subiectivă are o influență cauzală asupra acțiunilor comportamentale, a corpului material și a lumii exterioare.*

- *Qualia: Calități, mintea experimentează subiectiv atribute precum culorile, durerile, sunetele și așa mai departe*.

Aceştia susţin, de asemenea, că conştiinţa a evoluat în moduri distincte de-a lungul timpului, rezultând o diversitate considerabilă în ceea ce priveşte modul în care a fost instanţiată. Ei susţin că conştiinţa primatelor, în special a omului, este fundamental diferită de cea a altor specii. Aceste alte specii, mai puţin complexe, în virtutea propriilor caracteristici neuronale şi comportamentale, au forme de conştiinţă diferite, deşi la fel de valide ontologic.

Privind înapoi prin regnul animal, analiza lor îi determină să includă mamiferele, păsările, peştii şi alţi amfibieni, artropodele (crabi, insecte) şi cefalopodele ca având toate mecanisme neuronale, vieţi mentale şi repertorii comportamentale suficient de bogate pentru a se califica ca având conştiinţă ontologică. Ei observă că fiecare şi-a atins propriul nivel de reprezentare subiectivă folosind sistemele neuronale care au evoluat în familia filogenetică respectivă. Dar, susţin ei, toate au ajuns la un punct final similar: o minte cu proprietăţi fenomenale care reflectă cele patru caracteristici criteriale enumerate ale conştiinţei. Momentul la care analiza lor conduce este perioada cambriană[53], care a început acum aproximativ 540 de milioane de ani. Având în vedere că viaţa a apărut pentru prima dată pe planeta noastră în urmă cu aproximativ 3,7 miliarde de ani, Feinberg şi Mallatt presupun că a fost nevoie de aproximativ 3 miliarde de ani pentru ca conştiinţa să apară din bacterii şi alte specii preneurale pe care ei le presupun lipsite de sensibilitate.

Modelul Feinberg-Mallatt (care este dezvoltat mai detaliat în cartea lor recentă, *Consciousness Demystified*) se bazează pe presupunerea că conştiinţa necesită un sistem neuronal, iar criteriile pe care le identifică sunt etichetate drept „neuro-ontologice”. Un mic joc interesant de jucat aici este să ne întoarcem la criteriile lui Feinberg şi Mallatt pentru conştiinţă, dar să înlocuim „neuroni” cu „receptori”. Ceea ce avem acum este un set de caracteristici pe care le prezintă speciile unicelulare, caracteristici pe care le-am văzut în descrierea lui Jennings a reacţiilor Stentofs la expunerea repetată la un stimul nociv discutată în capitolul 2 şi pe care le vom examina în detaliu în capitolul următor.

Dacă plasăm originile sensibilităţii odată cu prima apariţie a vieţii, perioada de 3,7 miliarde de ani de la momentul în care s-au dezvoltat aceste prime forme de viaţă poate fi marcată nu de apariţia conştiinţei, ci de evoluţia unor noi forme, mediate printr-un cadru multicelular bazat pe procesele darwiniene, simbiogenetice şi epigenetice care au permis evoluţia ţesutului nervos. Feinberg şi Mallatt susţin că diversitatea în viaţa mentală pe care o citează a derivat dintr-o serie de apariţii distincte şi perceptibile ale conştiinţei (la fel cum zborul a apărut într-o varietate de specii, fiecare folosind o cale evolutivă diferită). De asemenea, este rezonabil să considerăm aceste varietăţi de conştiinţă nu ca tipuri distincte de experienţă subiectivă, ci ca simboluri, variaţii pe o temă care a fost stabilită atunci când au apărut primele procariote sensibile.

Feinberg şi Mallatt observă că modelul lor, bazat pe diversitatea sistemelor sensibile, necesită o serie de momente emergente distincte, fiecare dintre acestea creând propriul decalaj explicativ distinct între structurile materiale ale gazdelor de creiere diferite şi expresiile mentale fenomenale aferente. Din nou, nu văd niciun motiv să mă îndoiesc că aceste experienţe subiective distincte au loc în aceste diferite *phyla*. Eu doar pun la îndoială cadrul general pe care l-au creat pentru a le vedea. Occam ne-ar încuraja să le tratăm mai degrabă ca variaţii noi pe o temă decât ca tipuri ontologic distincte, fiecare dintre acestea necesitând propriul său mic miracol emergent.

Colin Klein şi Andrew Barron.

53 Primele aproximativ 20 de milioane de ani din această epocă se numesc „explozia cambriană” şi au fost martorii apariţiei celor mai multe filii moderne şi au pregătit terenul pentru diversificarea rapidă a speciilor care a avut loc în următorii 70-80 de milioane de ani. Înainte de perioada cambriană, practic toate formele de viaţă erau organisme unicelulare care trăiau în cea mai mare parte ca indivizi distincţi, cu câteva care au început să formeze colonii.

Klein, un filosof, și Barron, un biolog, au colaborat la o abordare multidisciplinară care îi conduce la concluzia că primele dovezi ale conștiinței se găsesc la insecte. Aceștia abordează problema pe aceeași cale metodologică ca Feinberg și Mallatt, adică identifică sistemele neuronale care stau la baza conștiinței umane și caută să vadă unde au apărut pentru prima dată aceste sisteme sau analogii clare ale acestora.

Citând dovezi din linii convergente de cercetare privind stările vegetative, leziunile cerebrale, anestezia și alte tulburări neuroclinice, aceștia identifică creierul mijlociu integrat și ganglionii bazali ca fiind zonele cerebrale care susțin toate formele de conștiință la vertebrate. Ei observă, de asemenea, că aceste structuri subcorticale procesează informații, având în același timp un impact redus (zero?) asupra conștientizării naturii informațiilor respective. Pe scurt, acestea sunt neurocorelațiile procesării implicite și susțin ceea ce am observat mai devreme că dă naștere unei distincții „**care** versus **ce**”. Adică, starea mentală în care o persoană știe că știe, dar este introspectiv oarbă la ceea ce știe. După cum am menționat, multe dintre experiențele noastre subiective sunt de acest fel - știm că am învățat ceva, dar adesea nu suntem conștienți de detalii. Regiunile subcorticale asupra cărora se concentrează Klein și Barron au fost, de fapt, identificate ca funcționând în acele stări de tipul „**știu asta, dar nu știu ce**”. Vom reveni asupra acestei distincții mai târziu.

Aceste sisteme permit oamenilor să realizeze ceea ce Klein și Barron consideră a fi caracteristicile critice ale conștiinței, și anume o integrare continuă a informațiilor din intrările perceptive, menținerea accesului la sistemele memoriale care ne mențin în contact cu cunoștințele relevante și interoceptorii care ne informează despre ceea ce fac corpurile noastre și permit monitorizarea feedback-ului motor.

Atunci când cercetează panoplia filogenetică, ei constată că insectele, deși nu dispun de structurile neuroanatomice specifice primatelor, au structuri analoge care îndeplinesc, în esență, aceleași sarcini. Ei stabilesc paralele observând că ceea ce se numește complexul central la insecte procesează informațiile spațiale astfel încât acestea sunt integrate cu mișcarea corporală și cu amintirile experiențelor trecute în moduri care reflectă modul în care acestea sunt tratate de zonele creierului mijlociu și de ganglionii bazali ai speciilor care au evoluat mai recent.[54] Aceste funcții și structuri analoge sunt, susțin ei, suficient de convingătoare pentru a concluziona că insectele experimentează o conștiință autentică. Ei observă că aceste caracteristici neuroarhitecturale cheie se găsesc și la vertebratele antice, inclusiv la *lamprede* și la alte specii asemănătoare peștilor, ceea ce îi determină să fie de acord cu Feinberg și Mallatt că conștiința autentică a apărut pentru prima dată în timpul exploziei cambriene (detalii în Klein & Barron. 2016 și Barron & Klein. 2016).

Peter Godfrey-Smith.

Cartea sa din 2016, despre care am vorbit mai devreme, a fost subintitulată *Octopus, the Sea, and the Deep Origins of Consciousness* (Caracatița, marea și originile profunde ale conștiinței), dar ar fi putut foarte bine să se numească *Ode to an Octopus*. În această carte, Godfrey-Smith combină observațiile din numeroasele sale ore de scufundări în largul recifelor din Australia, descoperirile științifice ale biologilor marini, biologilor celulari și anatomiștilor și, în cele din urmă, acordă cefalopodelor, în special caracatițelor și sepiilor, ceea ce el consideră a fi o minte autentică cu conștiință ontologică.

Godfrey-Smith remarcă faptul că, atunci când vine vorba de problemele legate de corp și minte, caracatița este o fiară foarte diferită, aproape extraterestră. Are un corp, desigur, dar este, în cuvintele sale, „un corp neînsuflețit”, deoarece nu are o formă fixă, nu are o formă continuă, ci este în mod constant modelat și reformat. În consecință, caracatițele au dezvoltat sisteme nervoase foarte diferite care, spre deosebire de cele din aproape toate celelalte specii, nu sunt centrate într-un singur loc, precum un cap sau o *notocordă*. Mai degrabă, după cum am observat mai devreme, ele au un

54 După cum s-a menționat, aceste zone, în mod curios, au fost, de asemenea, implicate în învățarea implicită, inconștientă la om. A se vedea P. T. Reber (2013) pentru o prezentare generală a acestei cercetări, cea mai mare parte a căreia se bazează pe tehnici fMRI (imagistică prin rezonanță magnetică funcțională) pentru identificarea structurilor cerebrale care sunt active în timp ce participanții se angajează în anumite acte cognitive.

număr de structuri neuronale asemănătoare creierului asociate cu fiecare braţ şi un creier central în capul animalului.

Godfrey-Smith subliniază aceste diferenţe în distribuţia neuronală, tipul de corp şi mediul pentru a susţine că conştiinţa şi mintea au evoluat în circumstanţe foarte diferite şi în specii foarte diverse. Ceea ce contează în abordarea sa sunt funcţiile cognitive ale speciilor, în special capacitatea de conştientizare şi autoreferenţiere. Concluzia sa este că există foarte mult „ceva ce înseamnă să fii o caracatiţă”.

Multe altele.

Într-adevăr, nu lipsesc alte sugestii şi multe puncte din filogenie în care ne putem aştepta să vedem cum apare conştiinţa sau conştientizarea ontologică reală. Sir John Eccles a stabilit acest punct în urmă cu aproximativ 200 de milioane de ani, când creierul mamiferelor a dezvoltat pentru prima dată un cortex cerebral primitiv. Dar Eccles a fost, de asemenea, intrigat de performanţele cognitive ale mai multor specii de păsări şi a considerat că, poate, *wulstul aviar*, un analog al izocortexului mamiferelor (denumit în prezent, în general, *hiperpallium*), ar putea fi responsabil - caz în care, în cadrul său, păsărilor li s-ar acorda conştiinţa. Pe de altă parte, reptilele au fost excluse, deoarece nu au un *hiperpallium*.

Irene Pepperberg, observând, de asemenea, că atunci când speciile aviare au evoluat din reptile, acestea au dezvoltat nu numai *hiperpalliumul*, ci şi, după cum s-a menţionat mai sus, seria de structuri cerebrale cunoscute în prezent ca fiind omologii neocortexului mamiferelor. În consecinţă, ea a plasat prima apariţie a conştiinţei la păsări, ceea ce, având în vedere lunga sa carieră în studiul funcţiilor cognitive ale papagalilor, nu este foarte surprinzător.

Elrik Sovik şi Clint Perry susţin că conştiinţa nu poate fi apărută înainte de evoluţia sistemelor nervoase şi identifică *bilateria*, grupul mare de animale cu simetrie bilaterală, inclusiv vertebratele, artropodele şi moluştele, ca fiind punctul de origine al experienţei fenomenale autentice - plasând din nou originile în explozia cambriană.

Bruno van Swinderen analizează cei 200 000 de neuroni din creierul muştei de fructe şi concluzionează că aceştia îi oferă suficientă funcţionare cerebrală pentru a manifesta atenţie selectivă, pe care o consideră funcţia necesară pentru o conştiinţă autentică.

Euan Macphail nu este de acord cu toate aceste sugestii. În analiza sa, limbajul generativ este necesar pentru apariţia conştiinţei autentice. Prin urmare, în opinia sa, conştiinţa a apărut mai întâi odată cu apariţia speciei care a dezvoltat prima limbajul, probabil unul dintre predecesorii noştri imediaţi.

În eforturile lor de a rezuma aceste diverse oferte, neurologul suedez Peter Arhem şi colegii săi B. I. B. Lindahl, Paul Manger şi Ann Butler au observat că problema profundă a tuturor acestora este că nu există un consens în domeniu cu privire la „markerul” comportamental, cognitiv sau biologic care ar conta ca semn empiric al conştiinţei ontologice. Aceşti oameni de ştiinţă fac, de fapt, doar presupuneri educate cu privire la ceea ce ei consideră a fi procesele sau mecanismele criteriale sau caracteristicile biologice suficiente pentru a face posesorul lor conştient. Atunci când au analizat fiecare dintre numeroasele scenarii propuse, ei au ajuns la concluzia că:

Fiecare [propunere], în sine [este] coerentă şi, în prezent, nu este posibil să se stabilească în mod concludent una dintre ele ca fiind cea mai probabilă. Toate scenariile sunt construite pe lanţuri de dovezi fragmentate, iar scenariul care va prevala depinde de momentul şi modul în care sunt contestate verigile cele mai slabe.

Ca şi în cazul modelelor care încearcă să descopere corelaţiile neuronale ale conştiinţei, aceste eforturi diferite de a identifica punctul din evoluţie în care apare pentru prima dată o minte autentică, o conştiinţă autentică nu vor rezolva *problema dificilă*. Ambele strategii de cercetare au dat rezultate ştiinţifice şi filozofice şi este cazul să le lăudăm pentru asta, dar Arhem şi colegii săi au

dreptate, până în prezent nimeni nu a prezentat decât „dovezi fragmentare”. În capitolul următor, voi încerca să remediez această problemă.

Câteva gânduri despre aceste eforturi

Înainte de a trece mai departe, nu mă pot abține să nu împărtășesc câteva gânduri pe care le-am avut în timp ce căutam acest material. Am fost surprins (și amuzat) de cât de des un cercetător părea să localizeze caracteristicile cheie ale conștiinței în specia pe care o studia sau în domeniul academic în care s-a specializat.

- Penrose, un fizician matematician cu o pregătire în teoria cuantică, a găsit rădăcinile conștiinței în mecanismele cuantice și, împreună cu Hameroff, un anestezist interesat de biologia celulară, a emis ipoteza că aceste mecanisme apar în microtubulii din celulele nervoase.
- Jaynes a petrecut mai mulți ani în Regatul Unit ca actor și dramaturg și a fost profund pătruns de clasicii antici – și a descoperit dovezile pentru modelul său în interpretările sale, după cum au spus unii, destul de neconvenționale ale scrierilor lui Homer.
- Macphail, un lingvist, plasează originile conștiinței în speciile cu limbaj generativ, adică *H. sap* sau versiunile proto-umane.
- Barron este specializat în neurobiologia furnicilor și albinelor și a reușit să ajungă la concluzia că originile conștiinței pot fi văzute în insecte.
- Biofilosoful Godfrey-Smith nu este doar un scafandru experimentat, ci și autor și coautor al unor articole publicate în reviste de biologie și în publicații populare despre biologia și comportamentul caracatițelor. Deloc surprinzător, el identifică cefalopodele ca un punct probabil pe marele arbore al evoluției pentru prima apariție a conștiinței.
- Dennis Bray, a cărui lucrare am discutat-o mai devreme, are o experiență de o viață în studiul organismelor unicelulare. În cele din urmă, el a concluzionat că sensibilitatea validă din punct de vedere ontologic este prezentă în speciile eucariote unicelulare și, poate, ar putea fi prezentă în procariotele mai simple.
- Gerald Edelman a descoperit că sistemul imunitar funcționează în conformitate cu principiile darwiniene și, câteva decenii mai târziu, a construit o teorie analogă a conștiinței umane bazată pe darwinismul neuronal.
- Todd Feinberg, un neurolog clinician, limitează conștiința autentică la speciile cu sisteme nervoase dezvoltate și funcționale.
- Sheets-Johnstone, dansatoare și coregrafă de formație înainte de a se îndrepta spre filozofie, consideră că cheia sensibilității se află în mișcare.
- Irene Pepperberg, o specialistă în cogniția aviară, cunoscută mai ales pentru munca sa asupra papagalilor, o găsește și la păsări.
- Klein și Barron, care au studiat creierele insectelor și au identificat analogii neuronale la structurile creierului uman, emit ipoteza că acestea sunt substraturile neuronale pentru originea sensibilității.
- Sovik și Perry sunt biologi care lucrează asupra bilateralismului și, prin urmare, se concentrează asupra speciilor bilateraliste ca punct probabil în dezvoltarea filogenetică pentru apariția conștiinței.
- Principalul subiect de cercetare al lui van Swinderen este *Drosophila melanogaster* - iar el și colegii săi identifică prima apariție a conștiinței în această muscă a fructelor, insecta preferată a geneticienilor.

Nu sunt sigur ce dezvăluie acest model, dacă este cazul, dar glumețul din mine se întreabă dacă panpsihicii (ei bine, panpsihicii darwiniști ai lui Hackel) nu au descoperit ceva. Indiferent unde te uiți, poți găsi dovezi sau un argument convingător pentru apariția minții. Iar locul în care cauți depinde în mare măsură de locul în care ai căutat tot timpul.

Dar, nu contează, ambele abordări strategice ale problemei conștiinței umane se confruntă cu dificultăți insurmontabile atunci când vine vorba de *problema dificilă*, de reducerea decalajului explicativ, de descoperirea modului în care materia generează viața mentală, în care fizicul poate genera experiențialul. Ambele sunt împovărate de două probleme profunde și irezolvabile: În primul rând, modul în care este exprimată însăși *problema dificilă* este rezultatul unei erori de categorie. În al doilea rând, toate suferă de dilema *emergentistului* - teoria **CBC** suferă de asemenea, dar, după cum am susținut, șansele de rezolvare cu succes a dilemei sunt mult mai mari în acest caz decât pentru toate celelalte abordări.

Întâlnirea cu o eroare de categorie

După cum am văzut, aproape toate programele contemporane de cercetare și abordările teoretice ale conștiinței încep prin a lua experiența fenomenală umană drept obiect principal de investigație. Această strategie, susțin eu, a fost ne-optimală deoarece, așa cum am sugerat în mai multe locuri, a condus la o eroare de categorie. De fapt, „duritatea” *problemei dificile* este tocmai rezultatul acestei poziții. Atunci când un filozof al minții sau un neurolog cognitiv repetă celebra întrebare „Cum face creierul mintea?”, creierul în cauză este aproape întotdeauna cel din interiorul craniului uman. Stevan Hamad, un neurocognitivist priceput din punct de vedere filosofic, a formulat întrebarea în felul următor:

[Î]n cazul în care organismele au minte, problema dificilă este de a explica cum și de ce astfel de mașini de supraviețuire/reproducere ar evolua mințile: Care este ***rolul*** *cauzal suplimentar și valoarea adaptivă a faptului de a avea o minte, în plus față de rolul cauzal și valoarea adaptivă a faptului de a avea doar capacitățile comportamentale în sine?* (Sublinierea în original.)

Atât timp cât astfel de întrebări sunt puse în această manieră, găsirea răspunsurilor va părea „dificilă”. Vinovatul aici este acel element adăugat, cel pe care Harnad îl suprapune peste acele capacități comportamentale mai primitive, pe care el - și alții, după cum am observat în discuțiile noastre despre durerea peștilor - le consideră ca fiind programe motorii lipsite de sensibilitate. Aceasta invită la o încadrare a problemei într-un mod care este profund problematic, deoarece presupune că a avea o minte adaugă ceva și că ceea ce a fost adăugat are un rol cauzal de jucat în crearea unui sine fenomenal despre care se presupunea că nu exista înainte. Aceasta este o mișcare complet inutilă. Toate organismele au minți, sau precursorii a ceea ce nouă, din perspectiva filosofiei minții, ne place să considerăm minți, deoarece toate sunt componente inerente ale formei organice. Nimic nu este „adăugat” în sensul că nu există afixe ontologic noi care să apară la un moment dat în evoluție. Este vorba de o acumulare treptată de funcții, toate având elemente sensibile, care sunt suprapuse și se întrepătrund cu cele preexistente.

Eroarea de categorie a fost una nevinovată și, cu siguranță, nu a început cu David Chalmers, deși el are meritul de a ne fi dat tuturor o mică sintagmă pe care nu mă pot abține să nu o scriu cu majuscule. Ea ne urmărește de foarte mult timp. De la protodualismul lui Platon, la legătura teologică dintre minte și suflet în perioada medievală, la reconfigurarea de către Descartes a problemei minte-corp, la filosofia contemporană și științele cognitive, totul a fost la un loc. Atunci când subiectul este mintea sau conștiința, se presupune că toate discuțiile viitoare se vor referi la o entitate recent evoluată care încapsulează un grup multifuncțional de operații care are loc la oameni sau poate, doar poate, la alte specii cu funcții cognitive cunoscute, relativ sofisticate. Alte forme de subiectivitate sau sensibilitate sunt excluse, uneori în mod explicit, ca în cazul lui Descartes, alteori în mod implicit, ca în cazul unei mari părți a ecologiei cognitive și a discursului filosofic recent.

Nu contează dacă anumite „alte” specii sunt sau nu incluse. Indiferent unde dorește un anumit teoretician să plaseze punctul de tăiere, pragul sub care funcțiile unei specii nu reușesc să obțină imprimaturul teoreticianului, aceeași eroare va fi comisă. De fapt, această mișcare, după cum tocmai am văzut, a dat naștere unei industrii artizanale în rândul biologilor, geneticienilor și

filosofilor, toți încercând să stabilească în ce punct al continuumului speciilor apare acest lucru al minții și străduindu-se să identifice proprietățile biologice, organizatorice care îl permit.

Din nou, permiteți-mi să fiu foarte clar. Nu am nicio problemă cu cercetarea empirică. Sunt un admirator de-o viață al lucrărilor din domeniul neuroștiințelor cognitive. Întotdeauna am considerat că principalele mele interese de cercetare se înscriu în această tradiție. Această activitate a oferit și va continua să genereze informații importante pentru înțelegerea creierelor, a tuturor tipurilor de creiere. Dar este foarte puțin probabil să rezolve **problema dificilă**. Cadrul **CBC** taie acest nod gordian conceptual. Toate experiențele sunt mentale. Toate organismele care simt, percep și se comportă au minte; toate au conștiință. Există o singură categorie de sensibilitate, deși există o mare multitudine de tipuri în cadrul acesteia, fiecare reprezentând tipurile particulare de experiențe fenomenale caracteristice speciei respective și care exprimă rădăcinile sale evolutive și nișa ecologică actuală.

Primul beneficiu al acestei repoziționări a problemei este că **problema dificilă** este înlocuită de alte probleme, care vor prezenta propriile dificultăți, dar cercetarea lor va fi un alt tip de întreprindere. În loc să încercăm să înțelegem neurocomplexitățile din creierul uman care dau naștere minților umane, accentul este redirecționat către probleme precum:

- Înțelegerea modului în care anumite tipuri de forme organice de bază, primitive, au ajuns să aibă biosensibilitatea care este fundamentul subiectivității
- Identificarea mecanismelor moleculare care au dat naștere țesutului nervos cu caracteristici bioelectrice, care au permis transmiterea mai rapidă a informațiilor și evoluția unor specii mai complexe și mai bogate în sentimente
- Descoperirea mecanismelor și proceselor complexe care au mediat evoluția speciilor pluricelulare Descoperirea mecanismelor care au permis speciilor pluricelulare să redistribuie funcțiile fenomenale de la celulele individuale la coliziuni ale acestora și, în cele din urmă, la funcții senzoriale, motorii, memoriale, cognitive și metacognitive

A începe prin a încerca să ne dăm seama cum creierele *H. sap*. produc mințile *H. sap*. a provocat numai probleme - cea mai confuză dintre toate fiind prezumția eronată că acestea, și tipurile de stări mentale pe care le produc, sunt cumva separate epistemic și distincte ontologic de alte forme de mentale și există pe un alt plan, de cealaltă parte a unui gol explicativ iluzoriu.

Desigur, alții au recunoscut problema. Eforturile lor au vizat determinarea locului în dezvoltarea evolutivă în care a apărut pentru prima dată o formă de minte care ar putea fi un simbol al aceluiași tip de conștiință pe care îl au oamenii. Acest lucru a început serios, după cum am menționat în capitolul 2, cu munca lui Donald Griffin și, după aproximativ un deceniu de rezistență, a devenit un obiectiv urmărit cu entuziasm. Problema, din păcate, este că atâta timp cât această împingere „în jos” nu coboară până la acel „turn al țestoaselor” până la originile vieții, adică până la cele mai simple specii unicelulare, și nu înțelege că sensibilitatea începe acolo, problema nu va fi rezolvată.

Urmarea acestei linii de argumentare are o consecință clară: un angajament față de un reducționism materialist, biologic.[55] După cum am menționat mai devreme, reducționismul are elemente nuanțate, iar unele dintre acestea le vom analiza mai târziu. Aici, vreau doar să observ că ceea ce propun este o strategie de cercetare comună în științele fizice. În fizica modernă a particulelor, de exemplu, materia este tratată ca având caracteristici care sunt aspecte inerente ale

55 Reducționismul este, în mod corect, un termen generic pentru un grup de principii științifice care sunt privite cel mai bine dintr-o perspectivă pragmatică. Cu alte cuvinte, exercițiile reducționiste nu sunt plonjări nebunești și necugetate către principii de bază ipotetice. Acestea sunt călătorii coerente efectuate cu un ochi la nivelul de analiză care oferă posibilitatea unei puteri explicative sporite pentru întrebările puse. Cu siguranță nu sunt singurul care susține acest punct de vedere. Alții au dezvoltat argumente similare vizând un reducționism psiho-neural, în special filosoful John Bickle.

formei sale fizice. Când aveți particule nucleare, obțineți *forțele nucleare slabă și puternică*; când aveți masă, obțineți *deformarea spațiului-timp pe care o numim gravitație*; când aveți electroni, obțineți *forța electromagnetică*. Asta este tot. Dacă se dovedește că există proprietăți mai fundamentale, dacă există un nivel mai profund, argumentul ar fi valabil și acolo. Pornim de la un Big Bang cu proprietățile noastre acum 13,8 miliarde de ani și ajungem la patru forțe: *nucleară slabă și puternică, electromagnetică și gravitațională*. Dacă viitoarea teorie fizică arată că toate acestea sunt produse de corzi sau brane multidimensionale care vibrează, va fi bine și așa.

Oamenii de știință din domeniul fizicii, în cea mai mare parte, nu au probleme cu acest tip de reducționism fundamental (a se vedea Carroll, 2016, pentru o prezentare generală elegantă a acestei probleme și a multor altora), dar mulți oameni de știință din domeniul social și filosofi au, iar acest lucru poate crea probleme. În mod tradițional, ne simțim atrași de complexitate, iar apoi trebuie să ne străduim să despachetăm numeroasele sale componente și să dezvăluim diferitele interacțiuni dintre acestea - evaluând în mod conștiincios dimensiunile efectelor și separând varianța. Adesea, această abordare funcționează. Acesta este modul în care am realizat unele dintre cele mai bune cercetări psihologice. Dar atunci când vine vorba de probleme precum conștiința și originile minții, aceasta a fost un impediment în calea progresului din cauza acestei erori de categorie. Și odată ce ați căzut în această capcană logică, nu puteți scăpa de următoarea: *emergentismul*.

Dilema emergentiștilor

Emergentismul a fost mult timp un fel de bubă în buzunarele filosofilor științei, precum și ale colegilor lor filosofi mai specializați care se concentrează pe minte și conștiință. Fiecare efect, funcție sau formă nouă care este studiată este văzută ca fiind rezultatul unei combinații complexe de operațiuni, obiecte și entități care determină cumva apariția unui efect nou, a unei proprietăți sau funcții noi. Adunați lumina soarelui, tipul potrivit de proteine și clorofilă și veți obține fotosinteza și produse noi precum zaharuri și oxigen. Combinați oxigenul și hidrogenul în modurile și cantitățile potrivite și apă apare cu proprietăți noi - este umedă și stinge incendiile pe care elementele sale constitutive le încurajează. Obțineți un creier cu cele aproximativ 100 de miliarde de neuroni fabulos de complicați, care interacționează între ei, cele aproximativ 100 de trilioane de sinapse și structurile și sistemele de sprijin asociate și veți obține... conștiința umană.

A afla cum se produc aceste miracole va fi întotdeauna o problemă. În conformitate cu standardele științei normale, răspunsurile trebuie extrase din suspecții obișnuiți: teorie, experimente, analiza datelor, argumente, un pic de postură și multă muncă grea. Dar, până să ajungem la cea despre creierul uman și conștiință, nimeni nu a crezut că ne confruntăm cu vreo criză epistemică. Fotosinteza a fost rezolvată. La fel și apa. Dar nu și conștiința. Atunci când conștiința este subiectul analizat la microscopul academicienilor, lucrurile par să devină întotdeauna complicate și, din nou, vinovată este eroarea de categorie. Caricaturistului Stanley Harris îi plăcea să îi ironizeze pe oamenii de știință, iar acest desen celebru din 1972 prezintă exact dilema *emergentiștilor.* A se vedea figura 3.1.

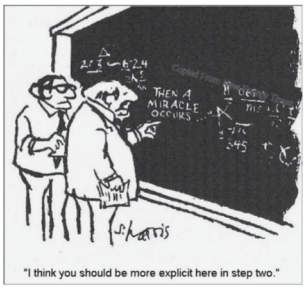

Figura 3.1. Perspectiva lui Sidney Harris asupra dilemei emergentiștilor. Retipărit cu permisiunea lui S. Harris la ScienceCartoonsPlus.com.

Dar, cu toată corectitudinea, deși îmi critic colegii, recunosc cât de ușor a fost/este să nu vedem ce se întâmplă. Deoarece conștiința, sensibilitatea, experiența fenomenală, subiectivitatea - oricum am dori să o numim - este atât de răspândită, atât de bogat manifestată în vasta panoplie a speciilor, mediilor și ecologiilor, de la adâncurile zdrobitoare ale faliilor din mijlocul oceanelor până la cele mai înalte platouri rarificate ale Anzilor și Himalayei, de la specie la specie este departe de a fi evident că toate sunt simple simboluri ale unui tip singular.

Această problemă a apariţiilor multiple ale unor forme particulare de conștiință este aproape inevitabilă. Specializarea abundă. Unele specii au vedere în culori, altele văd lumea prin ochi monocromatici, iar altele sunt practic insensibile la fotografie. Majoritatea speciilor au corpuri cu o conformație destul de rigidă; altele, cum ar fi multe cefalopode, sunt atât de fluide în această privință încât forma lor implicită este vizibilă doar atunci când sunt în repaus. Putem deține mai multe reprezentări mentale, putem descompune dependențe imbricate, putem face algebră, putem avea limbaj. Liliecii și cetaceele nu pot, dar pot identifica obiecte pe baza texturii ecourilor emise de semnalele auditive înalte autoproduse. Gama nesfârșită de diversitate a formelor mentale, afective și motorii de sensibilitate ne îndeamnă să ne întrebăm: avem de-a face cu forme cu adevărat distincte de conștiință? Sau ne confruntăm cu variații pe o temă?

Abordarea lui Antonio Damasio, așa cum am menționat mai devreme, ne invită să luăm în considerare faptul că acest lucru al sensibilității, această capacitate fenomenală, subiectivă, internă de reprezentare mentală ar fi putut evolua de mai multe ori. Sculptarea marmurei empirice de către Freidman și Mallatt spune o poveste similară. Analiza durerii realizată de Brian Key ne încurajează să ne gândim la experiențele speciilor de pești atunci când se confruntă cu un eveniment care dăunează țesuturilor ca având o ontologie fenomenală diferită de ceea ce ar experimenta un mamifer atunci când se confruntă cu același eveniment. Key ar putea susține că un pește are o emoție de tip Damasio care evocă o reacție instinctuală, dar, în lipsa structurilor corticale necesare, este lipsit de sentimente. Atunci când lumea înconjurătoare este bogată în diversitate, este dificil să observăm puncte comune.

Iată care este problema cu care se confruntă toate teoriile și modelele existente care se concentrează pe diversitate: Acele forme de viață mai puțin complexe decât cele în care teoreticienii au identificat apariția versiunii lor de conștiință autentică experimentează totuși cu adevărat ceva. Ele nu sunt recipiente inerte, goale din punct de vedere fenomenal. Orice organism viu este sensibil la evenimentele din lumea sa. Dacă aceste forme primitive de sensibilitate nu vor fi considerate ca îndeplinind criteriile care au fost stabilite ca adevărați indicatori ai conștiinței, atunci ce sunt ele? Cum clasificăm repertoriile lor comportamentale funcționale? Ce vom face cu faptul că formele și funcțiile lor sunt adaptative, flexibile în fața unui mediu în continuă schimbare, disponibile pentru mecanismele de selecție naturală, epigenetică și biosinteză pentru a crea alte forme noi, adaptative? O bacterie are câteva mii de situsuri senzoriale pe suprafața sa. Acestea sunt acolo pentru că oferă informații utile despre ceea ce se întâmplă în lumea acestui procariot.

Daniel Dennett, după cum am menționat mai devreme, preferă să considere aceste procese ca fiind reflecții ale unei competențe simțitoare care desfășoară comportamente adecvate fără niciun conținut mental - o mișcare care, după părerea mea, câștigă puțină putere explicativă deoarece, la fel ca toate aceste alte abordări, este, de asemenea, ridicată pe piedestalul *emergentismului*. Ce prag de complexitate neuronală sau de competență computațională poate identifica Dennett care, atunci când este depășit, dezvăluie brusc capacitatea de înțelegere? Și, ceea ce este și mai derutant, având în vedere că am identificat deja o jumătate de duzină de teoreticieni, fiecare dintre aceștia identificând un punct diferit în marea întindere a vieții în care susțin că a apărut adevărata conștiință, ne confruntăm cu o panoramă amețitoare de posibilități - niciuna dintre acestea nu are prea multe dovezi biologice moleculare care să o susțină.

Fiecare teorie a originii conștiinței trebuie să se confrunte cu dilema emergentiștilor. Fiecare va avea nevoie ca un miracol (sau poate mai multe) să se întâmple undeva, la un moment dat, la o anumită specie (sau cladă sau filo) din istoria evoluției. Acest lucru îi include pe panpsihici - fie că o recunosc sau nu - și include modelul meu **CBC**. Dar, și voi mai suna clopoțelul încă o dată, cadrul

CBC necesită cel mai mic şi mai uşor de rezolvat miracol, şi acesta trebuie să se fi întâmplat o singură dată. Odată ce sensibilitatea a apărut pe scenă - în momentul apariţiei vieţii, susţin eu - toate celelalte forme de experienţă mentală sunt mărturii ale acelui tip unic, variaţii pe o temă, la fel cum miriadele de forme de viaţă sunt mărturii ale unui eveniment unic petrecut în urmă cu aproximativ 3,7 miliarde de ani.

Anticiparea criticilor: Câteva argumente profilactice

Două poziţii adoptate de diverşi oameni de ştiinţă şi filosofi par să prezinte provocări puternice la adresa modelului *CBC*. Sunt sigur că mai există şi altele sau că vor apărea în curând, dar acestea sunt cele care au fost cel mai des ridicate în discuţiile mele cu colegii. Prima se bazează pe prezumţia că ceea ce eu numesc „dilemă" nu este cu nimic diferit de o serie de alte situaţii care au apărut în evoluţie. A doua se bazează pe o distincţie între evenimentele care sunt detectate sau simţite şi cele care sunt simţite sau experimentate.

Despre dileme.

Un critic rezonabil al modului în care am abordat *emergentismul* ar putea întreba: „De ce noţiunea de conştiinţă emergentă este o dilemă, în timp ce apariţia altor funcţii cognitive şi comportamentale bruşte şi poate surprinzătoare nu este?" La urma urmei, s-ar putea observa că tot felul de forme şi funcţii, abilităţi şi competenţe - zborul, camuflajul, vederea culorilor - au apărut din sisteme organice anterioare în care acestea nu erau prezente. De ce nu şi conştiinţa?

Răspunsul meu este simplu. Niciuna dintre aceste alte caracteristici nu este esenţială pentru viaţă. Toate au fost adaptări la anumite medii, biosfere, contexte ecologice. Acestea au fost evoluţii calitative care au oferit organismelor cu tipul potrivit de caracteristici într-o anumită nişă ecologică un avantaj selectiv, darwinist, în acel loc, la acel moment. Specia originală ar fi putut, şi în majoritatea cazurilor a făcut-o, să continue să supravieţuiască fără expresiile noi pe care le-au luat noile specii.

Conştiinţa, sensibilitatea, subiectivitatea sunt însă diferite. Această expresie reprezentaţională internă, această experienţă fenomenală este cu adevărat un tip ontologic distinct. Este esenţială pentru viaţa mobilă. Nicio specie care se deplasează liber într-un mediu în continuă schimbare nu are nicio şansă fără ea. Supravieţuirea ar fi imposibilă. Nu-mi pot imagina un set de circumstanţe în care ar fi putut apărea mai târziu decât la începuturile vieţii însăşi.

Despre senzaţie versus simţire.

Am abordat acest subiect mai devreme. Este argumentul pe care filosofii precum Dennett şi oamenii de ştiinţă cognitivi precum Oakley şi Halligan îl prezintă şi care susţine că toate aceste comportamente pe care le citez ca dovezi ale sensibilităţii ar putea fi doar operaţiuni mecanice, pur stimul-răspuns, codificate în genele speciei şi desfăşurate într-un vid fenomenal, lipsit de reprezentare subiectivă. Toate acestea ar putea fi doar instincte evoluate, asemănătoare celor ale roboţilor, care nu au nicio încărcătură epistemică. Neuroştiinţa cognitivă Stevan Hamad, în maniera sa succintă obişnuită, a ridicat problema făcând distincţia între un organism care simte ceva, dar nu simte. În invitaţia sa adresată participanţilor la un institut de vară pe tema „Problema celorlalte minţi", el a prezentat aceste aspecte ca fiind esenţiale:

Simţirea... nu este acelaşi lucru cu sensibilitatea; este sensibilitate dacă este simţită. Întrebarea fundamentală ... se referă la sensibilitate: ... care sunt procesele sau stările biologice

sunt simţite, mai degrabă decât doar executate (sau calculate)? Ce organisme biologice sunt capabile să aibă stări simţitoare? Care procese biologice sunt precursori sau componente ale sensibilităţii, dar nu sunt ele însele sensibile? Care ar putea fi avantajul adaptiv al proceselor senzitive faţă de cele nesenzitive? (Sublinierea în original.)

Problema mea cu această perspectivă este foarte simplă: Nu are sens şi cadrul conceptual pe care îl stabileşte, din punct de vedere ştiinţific, empiric, creează tot felul de dificultăţi. Atunci când un eveniment este perceput, acesta este simţit. Este experimentat. Este codificat ca o stare fenomenală subiectivă - chiar şi atunci când organismul care simte este unicelular. Altfel spus, această obiecţie ratează unul dintre punctele-cheie care diferenţiază teoria **CBC** de alte abordări. Da, aceste evenimente senzoriale ar fi putut fi (şi probabil că au fost) rezultatul unor bucle biologice de stimulare-răspuns, programate genetic, dar atunci când au avut loc, organismul a fost conştient de ele. Da, conştientizarea la o bacterie este un tip destul de primitiv de conştientizare, dar ea există, este simţită, experimentată, codificată, reprezentată.

Din nou, pentru că nu pot sublinia prea mult acest element, fără componenta internă, centrată pe sine, orice specie ale cărei acţiuni şi reacţii ar fi fost doar produsul final al unor operaţiuni mecanice necugetate ar fi ajuns în tomberonul darwinist. Un protozoar trebuie să aibă un motiv pentru a se îndrepta spre o sursă de nutriţie şi pentru a se îndepărta de o substanţă nocivă care l-ar putea afecta sau ucide. Motivul pentru care se comportă aşa este acela că are o experienţă senzorială inconfortabilă atunci când întâlneşte molecula care dăunează ţesuturilor şi una plăcută atunci când se confruntă cu nutrientul. Experienţa valorică, senzaţia de disconfort sau de agreabilitate trebuie să existe şi trebuie să fie simţită. Pentru a repeta, un procariot, în special unul care tocmai a încorporat câteva molecule de glucoză, este echivalentul unei cine la un restaurant de trei stele Michelin.

Să continuăm să insistăm pe această linie, deoarece are consecinţe empirice. Omida mea trebuie să cunoască diferenţa dintre frunzele hrănitoare ale plantei mele de busuioc şi seva aversivă a unei buruieni uriaşe. Experienţa subiectivă la întâlnirea cu fiecare este esenţială pentru supravieţuire. Nu poate fi o operaţiune mecanică, lipsită de conştiinţă, deoarece un astfel de proces lipsit de minte nu ar „şti” ce este o particulă bogată în nutrienţi sau cum să identifice o *buruiană* sau de ce nu ar trebui muşcată. Este ca în cazul roboţilor care joacă poker pe care i-am întâlnit în capitolul 1. Acestea sunt entităţi mecanice care pot învinge oamenii la versiunea *heads-up, no-limit* a jocului, în timp ce nu simt nimic, nu ştiu nimic despre poker - sau, de fapt, despre orice alt joc, sau, de fapt, despre orice altceva.

Orice organism viu care se confruntă cu o serie de stimuli şi evenimente diferite pe parcursul vieţii sale poate reacţiona la acestea într-un mod adaptiv numai dacă dispune de un sistem senzorial bogat în receptori care detectează aceste evenimente şi evocă o determinare afectivă, reprezentată intern, cu privire la natura lor. În plus, în virtutea sistemului său de reprezentare internă, acesta poate face faţă noutăţilor. Dacă este confruntat cu o toxină nouă care dăunează ţesuturilor, este pregătit să reacţioneze prin retragere, tocmai pentru că experimentează ceva profund aversiv. Poziţia Hamad-Dennett, care presupune că retragerea se bazează pe o „competenţă” înnăscută, lipsită de „înţelegere”, pur şi simplu (şi îmi cer scuze pentru că repet această mantră) nu are sens. Lumea este în schimbare şi niciun sistem programat nu poate supravieţui în ea. Doar unul care are o înţelegere internă, fenomenală, existenţială a valenţei evenimentelor pe care le întâlneşte. Trebuie să existe o sensibilitate ontologică la început, punctul de unde a început totul.

Mai mult decât atât, încercarea de a oferi o explicaţie ştiinţifică satisfăcătoare a acestor comportamente diferite folosind tipurile de procese mecanice, fără afecte, pe care Dennett, Hamad şi alţii le recomandă, este mult mai greoaie decât recunoaşterea subiectivităţii ontologice a fiecărui organism. Observaţi seria de întrebări pe care Harnad le cere participanţilor să le ia în considerare. Care procese contează? Ce organisme le au? Care sunt simpli precursori? Care au mai multă (sau mai puţină) valoare adaptativă? Adoptarea acestei poziţii nu reprezintă o abordare ştiinţifică eficientă a studiului conştiinţei şi nu este tocmai potrivită pentru o poziţie filozofică eficientă privind problemele minţii.

Pentru explicația senzației fără conștiință pe care o presupune Harnad, fiecare caracteristică a mediului trebuie privită ca și cum ar fi existat un număr mare de experiențe timpurii din punct de vedere evolutiv care au fost cumva codificate în ADN-ul speciei și i-au furnizat lanțul adecvat de acțiuni stimul-răspuns. Mai mult decât atât, fiecare trebuie privită ca un participant la un proces evolutiv continuu, astfel încât, la un moment dat, acest proces, care producea doar experiențe nesimțite, să poată produce brusc experiențe care sunt simțite - și ne confruntăm din nou cu o explozie de dileme *emergentiste*. Și, în cele din urmă, chiar dacă așa ceva ar face parte din trecutul îndepărtat al unei specii, tot nu ne-ar oferi o explicație satisfăcătoare a modului în care acea specie a făcut față unor circumstanțe noi, deoarece operațiunile mecanice, lipsite de inteligență, nu gestionează noutatea foarte bine sau chiar deloc. Adoptarea perspectivei **CBC** ne scutește de a mai avea de-a face cu acest grup de probleme care până acum s-au dovedit a fi de nerezolvat.

Rezumat

Suntem cam la jumătatea povestirii mele - partea ușoară. Am examinat o varietate de abordări care au fost adoptate pentru a descifra modul în care materialul creează mentalul, dacă artefactele pot realiza acest miracol, cum creierul poate cauza conștiința, cum ar putea apărea sensibilitatea din diferite tipuri de lucruri fizice. Le-am găsit pe toate insuficiente din mai multe motive. Dar unul dintre lucrurile pe care le-am învățat cu toții cei interesați de istoria științei este că identificarea defectelor modelelor sau teoriilor existente nu este deosebit de utilă decât dacă ai o alternativă respectabilă.

Pe măsură ce ne-am plimbat prin aceste abordări adesea inteligente și inovatoare, am semănat expunerea cu afirmații despre modelul **CBC**, ca și cum ar fi fost în mod clar adevărat și, evident, un cadru mai util în care să operăm. Așadar, este timpul să trecem la partea cea mai grea, să prezentăm suficiente dovezi pentru a susține că sensibilitatea începe odată cu prima apariție a vieții și că toate formele succesive de conștiință și tot felul de minți evoluate sunt descendenți evolutivi direcți. Pe scurt, trebuie prezentat un argument convingător că sensibilitatea identificabilă ontologic a apărut o singură dată și că orice alt proces epistemic este construit pe baza acesteia.

În plus față de încercarea de a-mi susține punctul de vedere în capitolul următor, vom vedea mai târziu (în capitolul 5) cum adoptarea poziției **CBC** are mai multe beneficii secundare. Schimbarea naturii întreprinderii ne va permite, de asemenea, să schimbăm modul în care sunt abordate o varietate de enigme filosofice, să oferim sugestii cu privire la modul de evitare a multor fundături epistemice pe care le-au întâlnit alte orientări și să oferim orientări pentru explorările empirice viitoare în microbiologie și știința evoluției. Așadar, haideți să facem o încercare!

4. Când devenim serioși: baza celulară a conștiinței

În primul rând, să începem prin a remarca faptul că strategia generală de cercetare pe care o susține abordarea *Bazei Celulare a Conștiinței* (**CBC**) nu este deloc nouă. Cadrul evolutiv din spatele acesteia este unul familiar: „începeți simplu”. Este o abordare care a fost adoptată de mulți în domenii de studiu din științele cognitive, cum ar fi memoria, percepția și emoțiile. Așa cum a spus Dennis Bray în încheierea cărții sale despre biologia celulară,

Cât de bine aș putea defini, dacă aș fi pus la treabă, percepția vizuală a unei portocale sau a unui urangutan într-un creier uman? Nu prea bine, cu siguranță nu la nivelul anumitor celule nervoase sau sinapse sau modele de semnale electrice. Dar ce s-ar întâmpla dacă aș întoarce lanterna asupra sa și aș folosi biologia celulară pentru a ilumina psihologia? Multe procese precum respirația și mișcarea au caracteristici comune atât la oameni, cât și la microorganisme. Ele au aceeași bază moleculară subiacentă și pot fi studiate în orice organism este mai convenabil, care se dovedește a fi adesea cel mai mic și cel mai simplu.

Euristica larg acceptată este căutarea rădăcinilor primitive ale funcțiilor complexe, căutarea acestora în spectrul speciilor și căutarea universalității. Ori de câte ori o funcție sau o formă se găsește pe scară largă, în special la nivel universal, puteți fi practic siguri că aveți în față ceva important. În plus, euristica de căutare a rădăcinilor comportamentelor și funcțiilor oferă aproape întotdeauna exemple relativ simple și mult mai puțin rezistente la sondarea secretelor lor funcționale decât formele mature observate la *H. sap*. Trei domenii de cercetare în care această strategie a dat rezultate vor servi drept exemple: modificările moleculare ale sinapselor care stau la baza memoriei, absorbția fotonilor de către pigmenții fotosensibili care funcționează în vedere și mecanismele biologice care controlează răspunsurile la stimulii nocivi și atractivi.

Puncte comune cu alte domenii de cercetare

Există multe domenii specifice în cadrul psihologiei și al neuroștiințelor cognitive în care înțelegerile de bază privind funcțiile umane au fost obținute privind înapoi în evoluție la rădăcinile lor din specii mai simple. O privire rapidă asupra a trei dintre acestea va fi utilă.

Memoria.

Lucrarea fundamentală a lui Eric Kandel privind biologia memoriei[56] a fost realizată pe *Aplysia calif ornica*, un melc de mare cu doar 20 000 de neuroni de dimensiuni convenabile, care s-au pretat la experimentele sale inteligente. Kandel era interesat de mult timp de învățare și memorie și a colaborat la o serie de studii privind electrofiziologia hipocampului[57]. Însă hipocampul, cu amplasarea sa incomodă și numeroasele sale legături cu alte căi și centre cerebrale, nu a fost deosebit de atractiv atunci când a ajuns să exploreze ceea ce el ajunsese să suspecteze: că modificările de bază ale sinapselor neuronilor în timpul învățării reprezintă probabil baza biomoleculară pentru formarea amintirilor permanente.

56 A se vedea, de exemplu, Castellucci, Pinsker, Kupfermann. & Kandel (1970).

57 Structură în formă de căluț de mare adâncită în fiecare dintre lobii temporali mediali ai creierului. Acesta joacă un rol esențial în formarea amintirilor pe termen lung și în reprezentarea spațială. Deteriorarea hipocampului duce la deficite grave de memorie și, în cazuri extreme, produce tipul de amnezie anterogradă totală observată la faimosul pacient H.M. (de la Henry Molaison, după cum s-a aflat după moartea sa în 2008). Molaison a suferit o rezecție a lobului medial-temporal pentru epilepsie intratabilă. Intervenția chirurgicală a fost un succes în sensul că epilepsia a fost atenuată, dar, după operație, el a fost complet incapabil să formeze noi amintiri pe termen lung, deși, ceea ce este important, memoria pe termen scurt și memoria implicită au rămas intacte. În mod interesant, recenta revizuire a datelor de către William Winter (Winter, 2018) sugerează că Molaison avea, de asemenea, patologie a lobului frontal.

Urmând exemplul altora, precum Alan Hodgkin, Andrew Huxley și Sir John Eccles, laureați ai premiului Nobel, care au descoperit acțiunile ionice ale transmiterii axonale studiind axonii gigantici ai calmarilor, Kandel s-a orientat către neuronii mari ai *Aplysia* și a descoperit modificările de la nivelul sinapsei care erau responsabile pentru stabilirea unor amintiri stabile. Ulterior, a primit Premiul Nobel pentru Fiziologie și Medicină pentru aceste cercetări. Este demn de remarcat faptul că a fost sfătuit de colegi să nu urmeze această cale pe motiv că este puțin probabil ca ceea ce se descoperă despre mecanismele neuronale la o specie atât de simplă să se generalizeze la oameni. Dar s-a întâmplat. Proteinele descoperite la *Aplysia* care provoacă modificări moleculare la nivelul sinapsei sunt aceleași care acționează și la noi.

Vederea.

O mare parte din ceea ce am învățat despre vedere, originile și semnificația sa evolutivă provine din cercetarea formelor primitive de sisteme vizuale întâlnite la o varietate de specii - inclusiv la organismele unicelulare. Se știe de mult timp că bacteriile reacționează la lumină. În 1982, Elena și John Spudich au descoperit că aceste forme primitive de fotosensibilitate, în special varietatea găsită în receptorii fototaxici ai unui procariot (*Halobacterium salinarum*), utilizează același biomecanism ca pigmenții vizuali din retina umană.

Alte probleme, cum ar fi deslușirea funcțiilor inițiale ale vederii, au fost abordate de biologii marini Anders Garm și Dan-Eric Nilsson în cercetările lor asupra ochilor primitivi compuși, fără lentile, găsiți pe vârful brațelor stelelor de mare. O serie de experimente i-au condus la concluzia că funcția principală a acestui sistem vizual de bază este de a direcționa mișcarea, în special mișcarea către medii sigure. Atunci când stelele de mare au fost îndepărtate de recifurile de corali preferate, ele și-au găsit rapid drumul înapoi. Stelele de mare orbite rătăceau fără țintă.

Afectivitatea.

Aceeași strategie generală a fost mult timp un element de bază al cercetării privind emoțiile. Psihologul Joseph LeDoux de la Universitatea din New York a remarcat, într-o analiză recentă a domeniului, cât de constant am rămas față de poziția expusă de Darwin în cartea sa din 1872, *Emoțiile la om și animale*. Studiul afectelor, susținea Darwin, este cel mai bine abordat pornind de la premisa că reacțiile emoționale sunt extrem de adaptative și că experiențele emoționale umane sunt extensii ale unor stări endogene similare întâlnite în spectrul speciilor - începând cu organismele unicelulare. În 1972, Robert Macnab și Daniel Koshland au demonstrat, într-adevăr, că bacteriile prezintă modificări sistematice ale mișcărilor de apropiere și evitare în funcție de concentrațiile de nutrienți și/sau toxine din mediul lor - dovezi clare că acestea au stări interne, subiective, afective, stări cu valențe, și le pot utiliza pentru a ghida acțiunile.[58]

Există multe alte exemple ale acestui tip de strategie de cercetare, dar cred că ideea este clară. Deși este cu siguranță adevărat că melcii de mare ai lui Kandel au doar forme limitate de reprezentare memorială, ei pot, ca și noi, să fie condiționați în mod clasic, iar mecanismele care susțin aceste funcții împărtășesc un nucleu neurobiologic comun cu al nostru. O bacterie sau o stea de mare pot sculpta spațiul vizual în moduri foarte diferite de modul în care o facem noi, dar putem găsi o legătură comună în implementarea pigmenților care absorb lumina pentru a furniza unei specii informațiile necesare pentru a-și naviga mediul. Viața emoțională a *Salmonellei* este destul de

58 Munca lui Macnab și Koshland dezvăluie, de asemenea, dovezi ale unui sistem de memorie la procariote, o problemă pe care o vom discuta mai în detaliu mai târziu.

simplă, dar toate indiciile arată că aceasta utilizează aceleași mecanisme ca vertebratele pentru a determina valența afectivă a obiectelor și evenimentelor.

Concluzia este că, așa cum înțelege orice biolog evoluționist, ceea ce funcționează în evoluție rareori se pierde cu totul. Acesta este un punct cheie; l-am remarcat anterior și vom continua această linie de gândire mai în detaliu mai târziu. Este, de asemenea, un principiu care a jucat un rol teoretic important în unele dintre lucrările mele anterioare privind biologia evolutivă a învățării implicite (inconștiente) (a se vedea Reber, 1989, 1993 pentru detalii) și are implicații analoge pentru abordarea **CBC** pe care o adoptăm aici. Pentru moment, apreciați că există motive darwiniste întemeiate pentru care aceste mecanisme primitive, dar adaptative, nu sunt abandonate, pentru care devin caracteristici stabile ale speciilor și pentru care servesc drept platforme pe care pot fi construite ulterior forme mai complexe, mai flexibile și mai sofisticate.

O scurtă paranteză despre un moment de nedumerire personală

M-am întrebat adesea de ce adoptăm în mod obișnuit această strategie de cercetare clar eficientă atunci când examinăm lucruri precum învățarea și memoria, percepția, sistemele motorii și emoțiile, dar am fost între ezitant și reticent în a o lua în considerare atunci când subiectul este conștiința sau mintea. La urma urmei, atunci când suntem lipsiți de noțiunile nematerialiste, mințile noastre sunt cu siguranță cele care percep, învață, stochează informații, reacționează, ghidează mișcările și evaluează experiențele. Așa cum am sugerat în repetate rânduri (ei bine, mai mult decât „sugerat” - a devenit o bătaie de tobe neîncetată), este vorba de acea tendință aparent irezistibilă de a ridica conștiința umană, mintea umană - de a o plasa pe un plan mai eteric, unul care este îndepărtat din punct de vedere conceptual, distanțat de mecanic, de fizic. Dacă, așa cum susțin unii filosofi și neuroștiințifici, conștiința umană nu se va supune modelelor mecaniciste, dacă neuroștiința nu poate umple golul explicativ, atunci îndepărtarea de fizicalism și reducționism își arată farmecul. Odată ce acest pas conceptual a fost făcut, aproape că devine o erezie să sugerăm să căutăm originile minții în procesele biologice ale speciilor simple.

În plus, există o reticență paralelă, dar la fel de derutantă, în rândul multor filosofi de a plasa conștiința, experiența, subiectivitatea în ceea ce alții consideră a fi un cadru evolutiv adecvat. Dennett a fost îngrijorat de această tendință ca un câine cu un os bine înțepenit, iar eu sunt profund de acord cu el. De fapt, filosofii care au luat apărarea dualismului, precum Chalmers, și cei care au îmbrățișat panpsihismul, precum Goff și Velmans, au respins în mod explicit o astfel de mișcare, susținând că nu va funcționa, deoarece astfel de eforturi se lovesc de vechea *nemesis*, dilema *emergentiștilor.* Dar pentru că cei care adoptă această linie nu pot înțelege cum tipurile noastre de experiență umană pot fi derivate din legi fizice și pentru că sunt împovărați de această lipsă de dorință de a merge până la capăt cu Darwin, ei s-au pictat singuri într-un colț conceptual și vor rămâne blocați acolo pentru mult timp - acest tip de vopsea metaforică nu se usucă.

Nu este surprinzător faptul că, atunci când filosofii vorbesc poetic despre experiență, atunci când se gândesc la minunile conștiinței, meditațiile se referă la experiențe complexe, umane, fenomenologice. Ei vorbesc despre dureri și furie, curcubee, eleganța percepută a Turnului Eiffel, melancolia regretului, simfonii compuse, poeți adorați și detestați. Este corect că o simplă experiență cromatică este un indicator al conștiinței la fel de convingător ca un coral de Bach, dar coralul de Bach este cel care captează imaginația filosofului. Nu sunt mulți cei care folosesc termenul „stare mentală” cu privire la funcționarea interioară a *wombats* sau a omizilor.

Totuși, din punctul de vedere al unui psiholog, nu este greu de înțeles cum un filosof poate ajunge în acest colț incomod din punct de vedere metafizic, înconjurat de vopsea metaforică udă.

Imaginați-vă-l pe Chalmers[59] încercând să explice una dintre ceea ce el numește „problemele ușoare” ale conștiinței, memoria. Imaginați-vă tipurile de enigme cognitive pe care le-ar găsi dacă ar aborda memoria în termeni de reminiscențe proustiene, filamentele evanescente ale episoadelor parțial reprimate din copilărie sau dispozitivele mnemotehnice răutăcioase pe care le folosesc studenții la medicină pentru a ține evidența nervilor cranieni. El ar putea ajunge să creadă că memoria este, într-adevăr, ceva care rezistă la orice fel de reducere fizicalistă. Dar am făcut multe progrese în ceea ce privește înțelegerea modului în care formele și funcțiile noastre minunat de complexe ale memoriei descind din forme și funcții fundamentale mai primitive și, din acest punct de vedere, putem face mult pentru a le demistifica.

Este timpul să schimbăm strategiile. Este timpul să schimbăm întrebările puse - și, așa cum s-a observat de multe ori, o astfel de schimbare în cercetare poate fi deschiderea către găsirea unor răspunsuri mai bune. Avem nevoie de o teorie reducționistă a sistemelor biologice și de o înțelegere a circumstanțelor în care o stare subiectivă însoțește exact tipul potrivit de stare biologică. Adevărata problemă nu este modul în care anumite tipuri de organizări complexe ale structurilor neuronale dau naștere vieții mentale umane, ci modul în care sistemele organice de complexitate variabilă permit evoluția treptată a sistemelor cognitive de bogăție variabilă și a sistemelor afective de subtilitate mai mare. Totul începe de la început, cu speciile unicelulare, pentru că, după cum a spus neurologul Dale Purves, aceste celule „sunt pietre de încercare pentru mentalitatea umană”.

O scurtă introducere cu privire la speciile unicelulare despre care vorbim

Până acum, am vorbit despre specii unicelulare precum *E. coli* și *Stentor* ca și cum acestea ar fi organismele originale care au apărut pentru prima dată în urmă cu aproximativ 3,7 miliarde de ani. Desigur, acest lucru nu este corect. Actualele bacterii, *archaea* și *eucariote unicelulare* care fac obiectul cercetărilor pe care le vom discuta în acest capitol sunt descendenții acelor specii originale. Au trecut prin propria lor evoluție. Nu există nicio modalitate de a determina cum au fost primele forme de viață, așa că facem singurul lucru pe care îl putem face. Vom presupune - așa cum trebuie să facă oricine privește această lume microbiologică - că mecanismele, procesele metabolice, comportamentele pe care le vedem la acești subiecți unicelulari sunt cam aceleași cu cele care au apărut pentru prima dată în supa biotică emergentă de atunci. În orice caz, aceasta nu este o presupunere nerezonabilă.

Din motive pe care nu trebuie să le detaliem aici (și care sunt doar parțial înțelese), mutațiile sunt mult mai puțin frecvente la speciile unicelulare decât la alte organisme mai complexe.

Pe măsură ce vom dezvolta imaginea funcționării actualei culturi unicelulare, va deveni clar că acest tip de continuitate structurală și comportamentală este aproape sigur corect, deoarece primele specii ar fi trebuit să facă față acelorași tipuri de contingențe de mediu ca și cele contemporane. Rețineți că, în timp ce următorii 3,7 miliarde de ani au condus la *Homo sapiens sapiens*, acei înaintași ai noștri, acele specii unicelulare rafinate, sunt încă foarte aproape de noi. Așa cum ne spune cu dragoste Ted Anton, observatorul literar al științei, în ultima sa creație, *Planet of Microbes* (Planeta microbilor),

Azi trăiesc, invizibili și aparent invincibili, în fiecare colț al planetei, de la izvoarele fierbinți din Yellowstone la vârfurile munților din Antarctica și până în interiorul corpului nostru - mai mult de o sută de trilioane. Totuși, nu vă speriați: mulți microbi sunt aliați în atingerea sănătății noastre - ca să nu mai spunem a planetei noastre.

59 Nu vreau să pară că mă iau de Chalmers. Este doar că el a articulat această poziție la fel de puternic și eficient ca oricine. El a numit problema dificilă „Problema dificilă”, așa că el este un reper evident.

Cadrul bazei celulare a conștiinței

Speciile unicelulare sunt saci de noroi extrem de complicați. Din nou, de la Bray,

Ele sunt populații eterogene de particule, mari și mici, imersate într-un bulion bogat de componente solubile. Ele sunt sisteme chimice extrem de complexe care conțin structuri moleculare de o sofisticare uimitoare, realizând nenumărate reacții în fiecare secundă.

Deloc surprinzător, le-a luat mult timp să evolueze. Prima apariție a vieții pe această planetă a avut loc în urmă cu aproximativ 3,7 miliarde de ani.[60] Pământul s-a format pe o perioadă de câteva miliarde de ani prin acreție gravitațională și bombardament din spațiu. Momentul în care a dobândit o stare relativ stabilă a fost undeva între 4,5 și 4 miliarde de ani în urmă. Prin urmare, de la momentul în care apa acoperea cea mai mare parte a planetei și până la apariția vieții a trecut un interval de aproximativ 500 de milioane până la un miliard de ani. În acest timp, așa-numitul Eon Hadean, bulgărele prebiotic din oceane a reușit să adune moleculele necesare pentru apariția primei specii unicelulare.

Aceasta este o perioadă foarte lungă de timp. Există o tendință de înțeles de a privi la ceea ce este/era necesar pentru viață și de a ne retrage consternați. Cum ar fi putut să se întâmple toate aceste lucruri, chiar și cu atât de mult timp la dispoziție? Așa cum spunea scriitorul de știință Robert Service într-un eseu publicat în revista *Science* în 2015,

Pentru ca viața să fi început, trebuie să fi existat o moleculă genetică - ceva de genul ADN sau ARN - capabilă să transmită schițele pentru producerea proteinelor, moleculele de bază ale vieții. Dar celulele moderne nu pot copia ADN și ARN fără ajutorul proteinelor. Pentru a face lucrurile și mai dificile, niciuna dintre aceste molecule nu își poate face treaba fără lipidele grase, care asigură membranele de care celulele au nevoie pentru a-și păstra conținutul în interior. Și, într-o altă complicație, enzimele pe bază de proteine (codificate de molecule genetice) sunt necesare pentru a sintetiza lipidele.

Dar, după cum subliniază Service în restul acestui articol, în ultimii ani au apărut o serie de idei în rândul celor care se numesc cercetători ai „originii vieții”. De exemplu, chimistul John Sutherland de la Universitatea Cambridge și colaboratorii săi au identificat recent o cale prin care ARN-ul s-ar fi putut forma fără a fi nevoie de enzime (a se vedea Patel etal. 2015 pentru detalii). Ei au reușit să creeze precursori ai acidului nucleic cu ajutorul unor molecule destul de banale, cum ar fi cianura de hidrogen și sulfura de hidrogen și puțină lumină ultravioletă.[61] Și este foarte probabil ca alte „scurtături” similare către viață să fie descoperite în anii următori.

Când obișnuiam să le expun acest gen de idei studenților mei, în special celor care nutreau noțiuni creaționiste, o replică obișnuită era să se întrebe cum ar fi putut apărea din întâmplare ceva atât de complex precum o celulă vie. Aceștia introduceau variații ale „argumentului ceasornicarului”[62] și indicau numărul mare de procese și mecanisme specifice care trebuiau să existe

60 Unii se gândesc că viața a apărut pe o cometă sau un asteroid care a aterizat pe Pământ cam în această perioadă. Această noțiune, oricât de interesantă ar fi, nu ajută. Nu face decât să trimită întrebarea cum a apărut viața într-un alt loc și în alt timp. Chiar dacă ar fi adevărat, principiile de bază ale cadrului CBC ar rămâne valabile. Aceste procariote invadatoare ar fi fost conștiente, cu minți extraterestre minuscule.

61 Sutherland și colegul său Matthew Powner de la University College London au împărțit premiul „*Origin of Life Challenge*” pentru această lucrare care examinează diferitele mecanisme care ar fi putut funcționa în timpul acestei lungi perioade prebiotice. Provocarea a fost lansată de chimistul american Harry Londsdale în 2011 pentru a încuraja cercetarea asupra celei mai fundamentale dintre probleme.

62 Argumentul că designul implică un designer, pe care creaționiștii încă îl mai scot ocazional din pivnița lor bine aprovizionată de falsuri pentru a argumenta împotriva evoluției, a fost bine și complet demascat (vezi Dawkins. 1986).

și prezentau cifre care sugerau că probabilitatea ca totul să se întâmple din întâmplare era extrem de mică, atât de mică încât practic nu putea fi un eveniment natural. Această poziție misterianistă era de obicei determinată de credințe religioase, de ignoranța acelor scurtături pe care le menționează Service și de faptul că este foarte greu să înțelegi cât de lung este un miliard de ani - dar odată ce înțelegi, improbabilul începe să pară mult mai puțin improbabil.

Extratereștrii ar fi putut să ne viziteze planeta atunci când s-au format mările, când fisurile adânci de pe fundul oceanelor emanau foc, când asteroizii plouau printr-o atmosferă înăbușitoare, iar apa proaspătă de ploaie care transporta molecule complexe de pe uscat curgea spre mări și își amesteca substanțele chimice cu cele care erupeau de dedesubt - și ar fi privit pânza lipsită de viață și s-ar fi gândit „aici nu se întâmplă nimic de interes biologic”. S-ar fi putut întoarce un milion de ani mai târziu și tot nu ar fi văzut nimic care să îi facă să creadă că, în cele din urmă, vom păși pe acest pământ, îl vom revendica, îl vom hrăni și, poate, dacă vom fi lăsați liberi, îl vom readuce la starea sa inițială lipsită de viață. Ar fi putut să ne viziteze din nou, încă un milion de ani mai târziu și apoi din nou, și din nou, aproape o mie din aceste vizite de un milion de ani înainte de a vedea prima înflorire a vieții autoreplicante. Și ar mai fi trebuit să treacă încă două miliarde de ani până la erupția masivă de specii pluricelulare a „exploziei” cambriene.

Există numeroase cercetări, teorii și speculații cu privire la modul exact în care a avut loc această tranziție de la simple molecule la celule vii, de ce se bazează pe carbon și apă, care procariote au apărut primele, bacteriile sau *arheii*, cum au evoluat eucariotele cu ADN-ul lor bine fixat într-o membrană internă încapsulată și care a fost ordinea fiecăruia dintre numeroșii pași necesari - ARN mai întâi? lipide mai întâi? proteine? Dar nu trebuie să ne preocupăm de aceste aspecte aici. După cum vom vedea, toate speciile unicelulare care au apărut în această perioadă au caracteristici comune. Toate s-au implicat în lumea înconjurătoare folosind strategii bio-comportamentale similare. Aceste trăsături comune stau la baza abordării **CBC**.

Primele forme de viață pentru care există dovezi solide în înregistrările fosile sunt covoarele microbiene. Acestea sunt compuse din mai multe straturi de bacterii ținute împreună de o peliculă formată din substanțe lipicioase pe care bacteriile le secretă. Deoarece pot ajunge să aibă o grosime de câțiva milimetri, au fost suficient de substanțiale pentru a deveni parte a înregistrărilor fosile, dar, cu siguranță, au fost precedate de primele organisme individuale care au alcătuit aceste colective. În formațiunile stâncoase din Quebec, Canada, există unele indicii ale existenței unor specii unicelulare individuale anterioare, dar nu toți microbiologii sunt convinși. Bacteriile individuale și alte procariote au trăit, desigur, în afara coloniilor, dar pentru a face acest lucru, trebuiau să existe procese specifice. Consensul în rândul biologilor celulari este că trei funcții critice trebuie să fie prezente pentru viață:

- *autoreplicarea, capacitatea de a crea altele de același fel*
- *Metabolismul, capacitatea de a absorbi substanțele nutritive, de a le prelucra, de a excreta deșeurile și de a repara leziunile celulare*
- *O membrană celulară externă care permite moleculelor corespunzătoare să se deplaseze înainte și înapoi prin ea, blocând în același timp alte substanțe care ar putea fi dăunătoare sau care nu au nicio valoare*

În cadrul abordării CBC este necesară o funcție suplimentară:

- *Sentimentul*

Deși această caracteristică epistemică nu face parte din modelul standard, este un element obligatoriu din perspectiva **CBC**. Pentru a avea un mod de autoreplicare de succes, mecanisme metabolice funcționale și un scut extern protector, trebuie să existe și un sistem subiectiv,

reprezentațional, care să fie sensibil la starea acestor procese și operațiuni și care să poată lua decizii și să se angajeze în comportamente adecvate pentru a menține starea lor funcțională.

Cu toate acestea, este probabil mai bine să ne gândim aici la sensibilitate ca la o co-consecință necesară a primelor trei caracteristici. Componenta mentală, subiectivă, nonmaterială este esențială. Pentru a avea un sistem care se autoreplică, se hrănește și se protejează, acesta trebuie să știe, într-un mod extrem de primitiv, că face aceste lucruri, cât de bine le face și ce schimbări și alegeri trebuie să facă pentru a continua să fie capabil să le facă, deoarece lumea se schimbă; este în flux și lipsa de reacție ar putea fi fatală. Acesta este punctul în care abordarea CBC ne spune să căutăm miracolul emergent al conștiinței. Aici materia a devenit minte și toate celelalte minți sunt rezultatul unor mecanisme evolutive care au funcționat de-a lungul unor perioade foarte lungi de timp.

Odată ce aceste proprietăți sunt la locul lor, avem viață. Pentru ca acest protouman (pentru că asta a fost) să aibă o șansă de a obține un rol, chiar rolul principal, pe scena evoluției, prietenul nostru unicelular va trebui să îndeplinească câteva sarcini esențiale:

> *- Locomoția, deplasarea pentru a găsi substanțe nutritive și pentru a scăpa de circumstanțe aversive*[63]
> *- să determine care dintre numeroasele obiecte din lumea sa sunt benefice și favorabile vieții și care nu sunt*
> *- Să stabilească care molecule din mediul înconjurător sunt nutrienți autentici, care sunt doar obiecte neutre, demne doar să fie ignorate și care sunt nocive sau toxice și ar putea provoca daune organice și cărora trebuie să li se interzică accesul sau să fie împiedicat contactul cu acestea*
> *- să detecteze caracteristicile mediului său pentru a obține informații relevante, inclusiv temperatura, lumina, forțele geomagnetice și gradienții de nutrienți astfel încât să poată lua decizii și să facă alegeri adecvate*
> *- să comunice cu alții pentru a controla ratele de diviziune celulară și pentru a ajusta reproducerea în funcție de nivelurile de nutrienți existente și de alte caracteristici ale mediului imediat. Acestea includ atât schimburi de informații între specii, cât și în interiorul speciilor*

Pe scurt, un organism unicelular trebuie să fie capabil să se miște, să perceapă lumea înconjurătoare, să aibă un simț al propriei forme fizice, să aibă o reprezentare a sinelui care să fie distinctă de simțul celorlalți, să fie conștient de diferența dintre lucrurile care sunt bune și promovează supraviețuirea și lucrurile care sunt rele și amenință existența, să ia decizii cu privire la cum, când și unde să se miște și să aleagă între opțiuni. Și trebuie să îndeplinească toate aceste funcții într-un cadru fizic care este în flux și în continuă schimbare. Sau, după cum a spus eminentul paleontolog George Gaylord Simpson în urmă cu peste 70 de ani,

Toate problemele esențiale ale organismelor vii sunt deja rezolvate în protozoarele ... unicelulare și sunt elaborate doar în om sau în alte animale pluricelulare.

O privire mai atentă la viața unicelulară

63 Unele specii unicelulare sunt sesile (de exemplu, Stentor, discutată anterior). Aceste specii unicelulare sesile sunt sensibile la curenții de apă și fac ajustări comportamentale adecvate pe baza informațiilor preluate de la moleculele transportate de maree și curenți. În anumite circumstanțe, o specie sesilă precum Stentor poate, după cum am menționat, să se desprindă de punctul de ancorare și să se mute într-un alt loc. Altele, cum ar fi bacteria Caulobacter crescentus, se divid asimetric, celula „mamă” rămânând fixată la locul său, în timp ce celula „fiică” înoată pentru a-și găsi propriul loc.

Până acum, am trecut doar pe lângă suprafață, abia am atins ceea ce aceste organisme simt, percep, cunosc și la ce răspund. Să aruncăm o privire mai profundă asupra unora dintre strămoșii noștri procarioți și eucarioți, în timp ce examinăm mai în detaliu de ce sunt ei capabili: Figura 4.1 prezintă mai multe vederi ale vechiului nostru prieten *Stentor roeseli*, Figura 4.2, o schiță care prezintă elementele esențiale ale unei bacterii tipice.

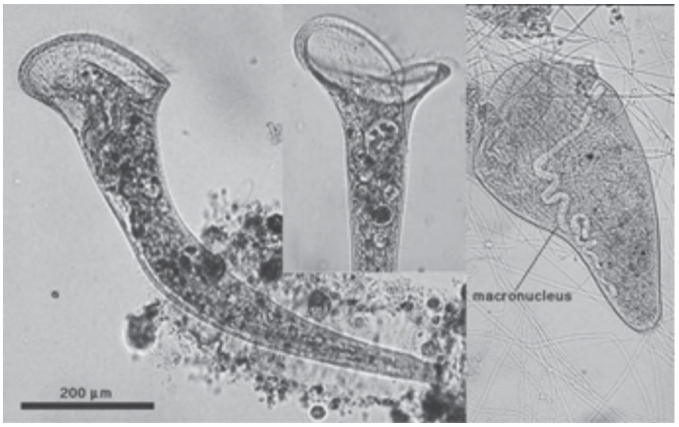

Figura 4.1. Stentor roeseli (compozit). Observați orificiul de admisie în formă de trompetă cu inele de cili pentru a aspira substanțele nutritive. Stentor sunt protiste și, ca și alte eucariote, au nucleu. Imagine în domeniul public (https://commons.wikimedia.org/w/index.php?curid=57073).

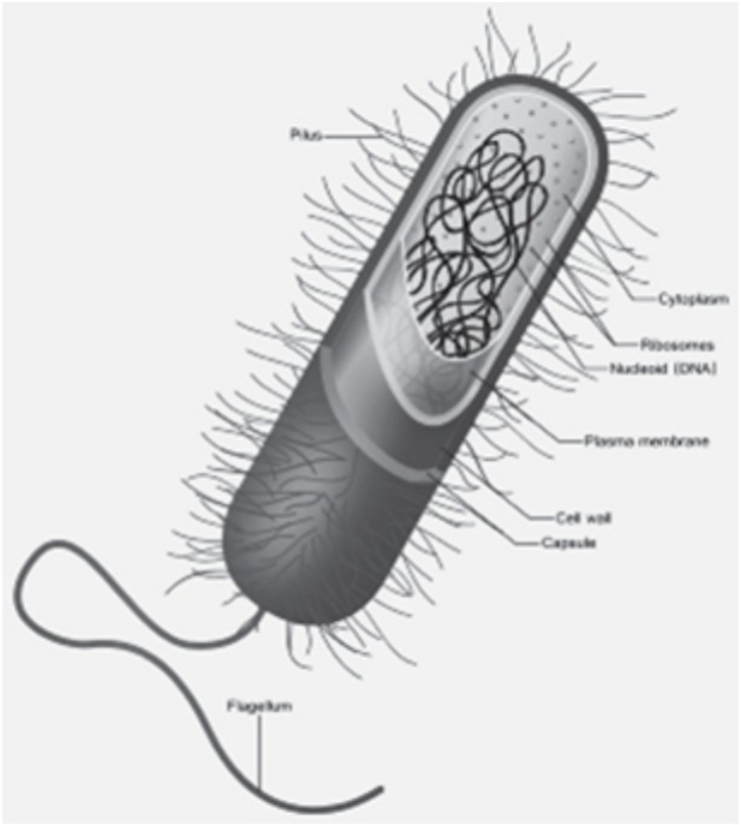

Figura 4.2. Schiță a unei bacterii procariote tipice care prezintă diferitele structuri și caracteristici. Rețineți că, spre deosebire de Stentor sau alte eucariote, ADN-ul nu este conținut într-un nucleu. Deși nu este prezentat aici, peretele celular este acoperit cu receptori chimiosensibili, așa cum se discută în text. Imagine de Ali Zifan. (https://commons.wikimedia.Org/wiki/User:OganessonOQ7).

După cum s-a menționat, elementele-cheie în stabilirea vieții necesită abilități senzorimotorii, capacități de luare a deciziilor și comunicare. Să analizăm mai îndeaproape modul în care aceste funcții funcționează în lumea celor foarte mici, folosind ca model *Stentor roeseli*, un protozoar de care am devenit destul de atașat. După cum am menționat în capitolul 2. *Stentor*:

- Folosește cilia de pe suprafața superioară pentru a ghida potențialele surse de hrană în pâlnia sa și în centrul celulei, unde le încorporează pe cele a căror structură moleculară indică faptul că sunt nutritive și le respinge pe cele care nu sunt. Această funcție este probabil cel mai elementar act pe care îl poate îndeplini, deoarece sistemul metabolic al organismului depinde de reușita de a distinge moleculele nutritive de cele care nu sunt nutritive.
- Detectează când moleculele absorbite sunt toxice și utilizează un răspuns vibrator eficient pentru a le expulza. Acest răspuns oarecum sofisticat are funcții adaptative clare. Deoarece Stentor se hrănește folosind cilia pentru a direcționa moleculele în jos, trebuie să fie capabil să le expulzeze dacă nu sunt nutritive. Alte specii unicelulare, precum Escherichia coli, detectează aceste proprietăți ale unei molecule atunci când aceasta intră în contact cu peretele celular și pur și simplu îi blochează intrarea sau se îndepărtează dacă nu este hrană.
- Învață să evite astfel de substanțe respingătoare după mai multe expuneri, retrăgându-și pâlnia în interiorul corpului său. Această capacitate este semnificativă deoarece dobândirea unui răspuns de evitare a fost considerată mult timp în psihologie drept o abilitate sofisticată. Evadarea este relativ simplă. Ea implică respingerea sau retragerea dintr-o situație sau stimul aversiv și este o reacție directă la un eveniment prezent (așa cum este exemplificat de răspunsul de respingere). Evitarea este mai complexă. Aceasta presupune recunoașterea consecințelor viitoare neplăcute ale stimulului respectiv, cu excepția cazului în care se adoptă un răspuns profilactic pentru a le preveni.
- Ia decizii calculate privind relocarea atunci când stabilește că mediul actual nu este sigur.

Toate aceste funcții se bazează pe stabilirea unor mecanisme de bază care implică abilități senzoriale, perceptive, motorii și decizionale și se regăsesc în întreaga lume a unicelularelor. Să analizăm în detaliu.

Funcțiile senzoriale și perceptive.

Sistemele senzoriale ale primelor procariote, cum ar fi bacteriile și *arhaea*[64], sunt foarte bine puse la punct, foarte variate și extrem de sensibile la mediul înconjurător al organismului. O bacterie tipică are aproximativ zece mii de chemoreceptori moleculari individuali, dispuși de-a lungul celulei. Sistemul chemotaxic al *Escherichia coli*, de exemplu, are cinci tipuri distincte de celule receptoare, ceea ce îi permite să detecteze aproximativ cincizeci de molecule de diferite substanțe chimice din mediu. Aceasta utilizează aceste informații pentru a direcționa mișcarea către sursele de hrană și alte circumstanțe favorabile și pentru a se îndepărta de substanțele care ar putea dăuna țesuturilor sau de alte substanțe aversive. Bacteriile sunt extrem de sensibile la gradienții moleculari și se vor deplasa rapid către substanțe atrăgătoare, cum ar fi zaharurile, și se vor retrage rapid din fața substanțelor respingătoare, cum ar fi acizii.

Alte specii sunt electrotaxice și folosesc un mecanism similar pentru a detecta particulele încărcate, pentru a compara stările interne cu condițiile externe și pentru a folosi diferențele pentru a se orienta către medii optime. Alte bacterii, precum *Magnetospirillum magnetotacticum*, sunt

64 Un grup separat de organisme unicelulare. Considerate odinioară ca făcând parte din grupul bacteriilor, acestea sunt acum clasificate ca un domeniu distinct. Cu toate acestea, ambele sunt procariote și nu au un nucleu încapsulat în membrană. Teoria general acceptată a apariției eucariotelor (toate celelalte specii la care ADN-ul se află în nucleu - inclusiv noi; noi suntem eucariote) este că prima a rezultat din încorporarea fortuită a unei archaea de către o bacterie - denumită în glumă Adam și Eva originali. Această ipoteză a fost avansată pentru prima dată la începutul anilor 1900 de botanistul rus Konstantin Mereschkowski și, ulterior, a fost dezvoltată în mod oficial de biologul Lynn Margulis ca parte a cercetărilor sale de o viață privind rolul simbiogenezei în evoluție.

magnetotaxice și utilizează informațiile geomagnetice împreună cu un mecanism aerotaxic pentru a găsi nu numai surse de hrană, ci și niveluri optime de oxigen. Și, poate că nu este surprinzător, bacteriile și *arhaea* prezintă fotosensibilitate și reacționează la lumină - deși ceea ce s-a dovedit a fi o mică surpriză când a fost descoperit pentru prima dată este faptul că folosesc rodopsina, aceeași moleculă de pigment fotosensibil ca și eucariotele, un domeniu care, desigur, include insectele, algele, vertebratele, păsările și mamiferele. Cititorii interesați pot consulta prezentarea generală din 2008 a lui John și Elena Spudich privind modul în care microorganismele utilizează ceea ce a ajuns să fie cunoscut sub numele de rodopsină bacteriană.

Într-o analiză a acestor diverse sisteme senzoriale și perceptive la bacterii, microbiologii Tobias Schweinitzer și Christine Josenhans au ajuns la concluzia că sistemele senzoriale bazate pe procesarea și modularea unei forme de taxis energetic sunt atât de utilizate de atât de multe specii diferite, încât le-au numit „globale”. Mecanismele care acționează aici sunt deosebit de subtile. Nu este suficient ca o bacterie să fie sensibilă la caracteristicile metabolice ale mediului înconjurător, ci trebuie, de asemenea, să-și monitorizeze propriile condiții metabolice interne și să compare cele două. Această abilitate comparativă, după cum spun ei,

este o condiție prealabilă pentru răspunsurile tactice la diferite niveluri de energie și randamente metabolice (permițându-le]... să navigheze către nișe metabolice mai favorabile și să supraviețuiască într-un anumit habitat.

Să ne oprim aici pentru o secundă și să punem o versiune a acelei „întrebări Dennett-Hamad”. Ne uităm la funcții pe care ar trebui să le numim „senzoriale” sau sunt mai bine considerate „perceptive”? În titlul acestei secțiuni, am folosit ambii termeni. Psihologii fac distincție între simpla simțire a unui stimul din mediu și perceperea acelui stimul, în cazul în care ultimul act este unul interpretativ. A simți căldura emanată de o tigaie fierbinte pe aragaz înseamnă a detecta faptul că în apropiere se află un obiect a cărui temperatură este considerabil mai ridicată decât cea a pielii tale. Perceperea faptului că tigaia este periculoasă și va provoca leziuni ale țesuturilor dacă este atinsă este un act mai sofisticat. A fi conștient de faptul că bucățica de mâncare pe care tocmai ați pus-o în gură este dulce este un eveniment senzorial; a ști că a fost o prăjitură cu ciocolată este un eveniment perceptiv. Senzația este simțirea; percepția este interpretarea. Descrierile făcute de Schweinitzer și Josenhans organismelor lor unicelulare reflectă acest set de funcții pe două niveluri, în sensul că este prezentă componenta interpretativă de „al doilea nivel” care diagnostichează percepția adevărată. O bacterie nu simte doar că există un gradient de molecule de zahăr în mediul înconjurător, ci percepe valența sa pozitivă și se deplasează activ în direcția concentrațiilor mai mari. O bacterie nu detectează pur și simplu faptul că o moleculă acidă îi atinge membrana de suprafață, ci interpretează acest lucru ca pe un stimul aversiv și se îndepărtează de el într-un mod clar măsurat.

În mod similar, biologii celulari Thorsten Mascher, John Helmann și Gottfried Unden au identificat mai multe mecanisme diferite în bacterii care se bazează pe „sisteme de transducție a semnalului”. Acestea funcționează prin *kinaze* (enzime catalitice) care se află pe membrana celulară, detectează (adică simt) stimulii de mediu și trec informațiile într-un sistem de reglementare internă care mediază răspunsul celulei în mod corespunzător (adică percepe). Sistemul de receptori este foarte variat deoarece, după cum s-a menționat anterior, bacteriile pot detecta până la 50 de molecule distincte. Sistemul de percepție are trei operații de transducție diferite, fiecare utilizează o tehnică funcțională ușor diferită care, în esență, oferă organismului informațiile cheie despre natura și valoarea stimulilor și îi permite să determine dacă să continue sau să se retragă. Mascher și colegii săi sunt destul de neechivoci cu privire la aceste mecanisme. Primul este denumit în mod regulat un sistem care „simte”, iar cel de-al doilea unul care „percepe”.

Pe scurt, avem de-a face cu organisme unicelulare care prezintă fără ambiguitate forme primitive, dar eficiente de sensibilitate.

<u>Mişcarea</u>.

Microorganismele dispun de o serie de tehnici pentru a se deplasa în mediul lor. Bacteriile se deplasează folosind mai multe metode diferite. Unele sunt împodobite cu mici filamente asemănătoare părului şi, în esenţă, se mişcă pe suprafeţe şi înoată prin lichide. *Amoebele* se târăsc practic prin extinderea unui pseudopod; restul celulei se îndreaptă spre extensie, mişcând celula înainte. *Amoebele* sunt, de asemenea, capabile să determine dacă un obiect întâlnit poate fi escaladat sau trebuie ocolit. Alte specii unicelulare de bacterii folosesc unul sau mai mulţi *flageli*, pe care îi pot roti ca nişte elici pentru a naviga prin apă. Unele, la fel ca multe eucariote cu *flageli* mult mai mari, îi folosesc pentru a propaga valuri de-a lungul lungimii lor pentru a le propulsa înainte, altele se rostogolesc în mod controlat, iar altele, precum *E. coli*, combină metodele de locomoţie şi înoată în mod ondulat şi, folosind mai mulţi *flageli*, se angajează în rostogoliri controlate. *E. coli* poate, de asemenea, să schimbe rotaţia *flagelilor* săi, folosind rotaţia în sensul acelor de ceasornic pentru a avansa şi în sensul invers acelor de ceasornic pentru a se retrage.

Mişcările bacteriene pot fi asimilate unor plimbări aleatorii „tendenţioase”, în care elementele tendenţioase intervin ca factori elementari în mediul înconjurător, instigând, împingând şi determinând mişcările - toate acestea fiind adaptate la informaţiile furnizate de sistemele lor perceptive. Imaginaţi-vă că mergeţi pe un drum. Detectaţi şi evitaţi obiecte, simţiţi mişcările persoanelor, automobilelor, bicicliştilor, animalelor. Identificaţi semnalele care vă spun să opriţi, să mergeţi, să viraţi. Nu a fost atât de diferit cu primele noastre specii unicelulare. Mediul era mai simplu, ceilalţi agenţi erau doar diferiţi compatrioţi unicelulari, obiectele erau simple molecule, unele comestibile, altele nu, semnalele erau destul de lipsite de ambiguitate, dar toate paralelele sunt acolo. Şi, la fel ca şi plimbarea noastră pe Main Street, primele organisme unicelulare au trebuit să simtă, să perceapă şi să analizeze natura lumii din jurul lor. Aşa a început totul şi o facem şi astăzi. Aşa cum a spus Sheets-Johnstone atunci când a contemplat organismele unicelulare,

[O] creatură ar putea cu greu să iniţieze o mişcare - să schimbe direcţia, să mărească viteza, să facă o pauză, să întindă mâna cu un apendice, să se întoarcă, să evite un obstacol sau un prădător, să exploreze, să fugă sau să se mişte intenţionat în nenumărate alte moduri [în timp ce]... este în concordanţă atât cu situaţia în care se află, cât şi cu propria sa corporalitate spaţio-temporală imediată.

În acest pasaj, ea se concentrează pe aceeaşi caracteristică cheie identificată mai sus. Capacitatea esenţială de a fi conştient de diferenţele dintre compoziţia lumii din jur şi cea din interior. Pentru a simţi prezenţa unei molecule de zahăr şi pentru a se mişca deliberat spre ea, o bacterie trebuie să ştie, la un anumit nivel, că aceasta este un obiect valoros, că starea sa metabolică internă este una care are nevoie de ea şi că anumite tipuri de acţiuni fizice îi vor permite să se mişte, să captureze şi să încorporeze molecula respectivă. Există aici o sensibilitate ontologică. Este greu de crezut să ne imaginăm aceste tipuri de acţiuni secvenţiale şi dirijate care se desfăşoară fără sentimentul de sine ca agent distinct cu un nivel palpabil de autonomie. În plus, eficacitatea lor este optimizată de faptul că sunt mobile. O celulă staţionară este la mila a tot ceea ce se întâmplă.

Motilitatea s-ar putea dovedi a fi mai mult decât o caracteristică esenţială a existenţei vieţii unicelulare. S-ar putea foarte bine să fie o caracteristică esenţială pentru menţinerea sensibilităţii, a subiectivităţii, a unei minţi conştiente. În capitolul următor, vom aborda o problemă care a planat în culise, în afara reflectoarelor: plantele, lumea florei. Pentru a anticipa problema de bază, plantele nu s-au ramificat în propria lor linie evolutivă până în urmă cu aproximativ 1,6 miliarde de ani sau aproximativ 2 miliarde de ani după apariţia primelor procariote. Dacă, aşa cum susţinem aici, acele specii unicelulare iniţiale erau conştiente în modul descris, atunci ar părea logic ca plantele, atunci când au evoluat pe cont propriu, să fi purtat cu ele acest tip de subiectivitate experienţială. La urma

urmei, un principiu de bază pe care l-am abordat de mai multe ori și pe care ne vom baza în continuare în capitolul următor este acela că evoluția nu renunță în mod obișnuit la adaptările de succes. Ea le menține; conservarea funcției este regula generală. Dar, din cauza modului în care au evoluat pereții celulari ai plantelor și a eșecului de a dezvolta motilitatea, este posibil ca acesta să fi fost un caz în care o funcție utilă a fost abandonată. Și da, desigur, există două interpretări alternative evidente: (a) sentimentul presupus a fi prezent în bacterii nu a existat și poziția **CBC** este complet greșită sau (b) plantele sunt sensibile. Acestea sunt întrebări empirice, nu teoretice - iar acest tip de viziuni alternative ale realității este motivul pentru care știința este atât de distractivă.

Mai multe despre acest subiect în capitolul următor. Pentru moment, să analizăm o serie de alte funcții pe care le prezintă locuitorii lumii unicelulare pentru a vedea cât de sofisticate și complexe sunt experiențele lor.

<u>Învăţarea</u>.

Ați putea anticipa că, de la lucrările lui Jennings de la începutul anilor 1900, modelul standard pentru speciile unicelulare este că acestea sunt mai mult decât niște automate de reacție stimul-răspuns ale căror comportamente pot fi descrise în termeni mecanici simpli - sau, mai direct spus, că pot învăța. Dar impactul volumului său din 1906 s-a estompat în timp. Interesant este că, în ultimul deceniu, s-a redescoperit capacitatea de achiziție a bacteriilor și a altor specii unicelulare. Orientarea cercetării către bacterii, din punctul de vedere al modelului **CBC**, este un pas interesant, deoarece acestea sunt procariote, adică „înainte de nucleu”, în timp ce *Stentor* este un *eucariot* („nucleu adevărat”).

Eucariotele, chiar și cele unicelulare, au structuri biomecanice mult mai complexe și repertorii comportamentale mai complexe decât procariotele.

Lucrările recente ale mai multor grupuri de cercetare prezintă dovezi clare ale condiționării simple de tip pavlovian la bacterii. Un laborator de cercetare condus de Yitzhak Pilpel de la Institutul Weizmann din Israel a descoperit că atât *E. coli*, cât și o drojdie, *Saccharomyces cerevisiae*, pot învăța să „prevadă” (termenul lor) evenimente viitoare și să înceapă să răspundă înainte de declanșarea acestora (a se vedea Mitchell et al., 2009 pentru detalii). Bacteriile adoră zaharurile - de fapt, mai mult decât „adoră”, acestea sunt singura lor sursă de energie. În studiul lor, bacteriile au primit în mod repetat lactoză și, la scurt timp după aceea, un alt zahăr, maltoza. În timp, bacteriile au arătat o schimbare în răspunsul lor genetic la maltoză și au făcut ajustări anticipative în funcția metabolică pentru a procesa maltoza mai eficient. Pentru a demonstra că aceasta a fost o învățare autentică a unei perechi stimul-stimul, au schimbat ordinea și au prezentat mai întâi maltoza. Bacteriile nu au prezentat răspunsul genetic, ceea ce indică faptul că nu a fost vorba de o simplă sensibilizare (care este o formă de învățare, dar nu una foarte interesantă), ci de dobândirea reală de cunoștințe despre o pereche ordonată.

Pentru a închide cercul conceptual pavlovian, Pilpel și colegii săi au arătat că răspunsul condiționat poate fi stins. După mai multe prezentări de lactoză singură, fără următoarea doză de maltoză, bacteriile au încetat să afișeze răspunsul genetic la maltoză - la fel cum câinii lui Pavlov au încetat răspunsul salivar la clopoțel după mai multe sunete fără mâncare.

O condiționare anticipativă similară a fost descoperită de Saeed Tavazoie și colegii săi de la Universitatea Princeton (pentru detalii, a se vedea Tagkopoulos. Liu. & Tavazoie, 2008). În loc să utilizeze proceduri directe de asociere, grupul lui Tavazoie a profitat de tendința naturală a *E. coli* de a asocia temperaturile ridicate (în jur de 37° C sau aproximativ 100° F) cu condițiile anaerobe ca, de exemplu, în intestinul mamiferelor. Coloniile de bacterii au fost apoi crescute în vase de laborator care au schimbat temperatura de la relativ rece la destul de cald și au ajustat nivelurile de oxigen de la niveluri ridicate la aproape deloc. Aceste condiții corespund, în linii mari, temperaturii și nivelului de oxigen din gura și intestinul mamiferelor. Având în vedere rolul jucat de *E. coli* în digestie, puteți înțelege de ce au fost alese aceste temperaturi și niveluri de oxigen.

Pe parcursul mai multor generații, ei au observat același tip de răspuns genetic pe care l-a observat grupul lui Pilpel. Colonia a învățat să anticipeze viitoarele schimbări de temperatură și de nivel de oxigen prin modificarea activității metabolice corespunzătoare circumstanțelor viitoare și, în mod interesant, a fost capabilă să inverseze asocierea naturală dintre temperaturi mai ridicate și condiții anaerobe. Aceste rezultate sunt importante deoarece implică coordonarea multor celule individuale într-un colectiv și pe parcursul mai multor generații. Este considerat un efect generațional deoarece bacteriile, în condiții optime de temperatură și nivel de nutrienți, se pot reproduce la fiecare 20 de minute. După cum am menționat în mai multe rânduri, aceasta este o problemă critică, deoarece va trebui să înțelegem mecanismele care intervin atunci când se formează colectivități, iar capacitățile senzoriale ale organismelor individuale sunt subsumate subiectivității globale a unui organism pluricelular.

Examinând alte specii unicelulare, Audrey Dussutour și colegii săi de la Universitatea din Toulouse au demonstrat recent că învățarea se manifestă la mucegaiul unicelular protist[65] *Physarum polycephalum*. *P. polycephalum* nu este deosebit de atașat nici de chinină, nici de cofeină. Dussutour și echipa sa au creat o punte între două plăci Petri cu nutrienți dezirabili și au împletit puntea cu pete de cofeină și chinină (a se vedea Boisseau, Vogel și Dussutour. 2016 pentru detalii metodologice). Pe parcursul mai multor zile, mucegaiul a devenit mai priceput la navigarea pe pod, evitând în mod sistematic locurile despre care aflase că erau amestecate cu substanțele amare. Grupul lui Dussutour părea aproape surprins de descoperirile sale, poate neștiind de unele dintre celelalte rezultate pe care le-am acoperit, notând că:

Rezultatele noastre evidențiază diversitatea organismelor lipsite de neuroni, care prezintă probabil o capacitate de învățare nerecunoscută până în prezent----Învățarea la organismele fără neuroni, cum ar fi mucegaiurile de nămol, este extrem de importantă pentru o înțelegere filogenetică cuprinzătoare a momentului și locului în arborele vieții în care au evoluat cele mai timpurii manifestări ale învățării.

Memoria.

Dacă microbii dau dovadă de învățare asociativă și fac schimbări adaptative în comportament și metabolism, implicația naturală este că ei au o modalitate de a instanția noile informații și de a le păstra. Pe scurt, pentru a avea o învățare autentică care se manifestă în timp, trebuie să existe o formă de memorie.

De fapt, se știe de ceva timp că organismele unicelulare formează amintiri. După cum am arătat mai devreme, Robert Macnab și Daniel Koshland au publicat dovezi convingătoare pentru un sistem memorial în bacterii în 1972. Interesant este că limbajul folosit de Macnab și Koshland în acea lucrare reflectă ceea ce bănuiesc că a fost un vag sentiment de neliniște în legătură cu găsirea de dovezi pentru o funcție cognitivă atât de evidentă la o specie unicelulară. Deși descrierea comportamentului bacteriilor pe care le-au studiat sugera cu tărie un sistem intern de reprezentare care deținea informații despre evenimente care au avut loc în trecut, ori de câte ori apărea cuvântul memorie, acesta era îmbrăcat în ghilimele ca „memorie” sau calificat ca „asemănător memoriei”. După cum s-a menționat anterior, anii 1970 au fost perioada în care etologul cognitiv Donald Griffin a fost atacat pentru că a sugerat că multe dintre abilitățile cognitive despre care se credea că sunt prezente doar la oameni erau, de fapt, prezente la o serie de alte specii multicelulare, „macro”. Cu siguranță, putem fi înțelegători cu ezitarea lui Macnab și Koshland atunci când relatează despre procariote.

Din motive care sunt departe de a fi evidente, descoperirea lui Macnab și Koshland și multe altele care i-au urmat au rămas sub radarul neuroștiințelor cognitive și, ceea ce este important pentru temele acestei cărți și ale modelului **CBC**, în mare parte nerecunoscute de cei care lucrează în

65 Protiștii includ orice eucariot care nu este un animal, o plantă sau o ciupercă.

filosofia minții. Cu toate acestea, sunt optimist. Există motive întemeiate pentru a crede că are loc o schimbare de orientare. Mai multe laboratoare de cercetare examinează în mod activ memoria la procariote, numind-o pe adevăratul ei nume, iar rezultatele lor atrag atenția.

Mai multe publicații mass-media din presa populară au publicat, de asemenea, articole axate pe fenomenul general. Scriitorul științific Jason Tetro a publicat un articol intitulat *The Mechanism of Microbial Memory* (Mecanismul memoriei microbiene) în numărul din 3 martie 2015 al revistei *Popular Science*. Site-ul de internet *Science Daily* a publicat un articol despre descoperirea recentă că bacteriile care fac parte din colectivități prezintă un tip neobișnuit de capacitate de memorie „de grup”. Sunt încurajat de această mișcare de recunoaștere a abilităților epistemice remarcabile ale acestor specii unicelulare și de descoperirile și mai intrigante cu privire la bio-mecanismele responsabile.

Baza organică a memoriei în cazul speciilor unicelulare.

Mai multe laboratoare de microbiologie care studiază învățarea la specii precum *E. coli* au examinat baza organică a acestor amintiri simple și, așa cum se întâmplă adesea în aceste contexte, se dovedește a fi mai complexă și mai sofisticată decât s-a crezut inițial. Biologii celulari Josep Casadesus și Richard D'Ari, de la Universitatea din Sevilla, au observat că bacteriile individuale au mai multe mecanisme de memorie distincte, adică diferite modalități de a codifica evenimentele și de a păstra informațiile pentru utilizare ulterioară. Unele sunt genetice, cum ar fi cel în care o funcție ereditară este codificată în secvențe specifice de ADN ca răspuns la experiență. Acest mecanism funcționează prin modificarea ADN-ului organismului într-un alt model de expresie genetică pe baza expunerii la anumiți stimuli de mediu. Acesta a fost tipul de proces care stă la baza descoperirilor laboratoarelor Weizmann și Princeton pe care le-am descris mai devreme. După cum spun Casadesus și D'Ari, acest mecanism „fixează” o anumită stare, deoarece fiecare organism nou care rezultă în urma diviziunii celulare poartă cu el această memorie nouă, codificată genetic. Alte mecanisme se bazează pe modificări ale proceselor celulare care sunt rezultatul unor stimuli sau acțiuni repetate și nu sunt ereditare. La fel ca o mare parte din ceea ce se învață, cunoștințele mor odată cu individul care le-a dobândit.

Este rezonabil să ne întrebăm care ar putea fi avantajul adaptiv al memoriei în viața unei bacterii? Răspunsul este simplu și, odată ce te gândești la el, evident. Pentru a putea reacționa eficient la modificările concentrațiilor moleculare din mediul înconjurător, organismul individual trebuie să poată păstra o înregistrare a circumstanțelor din trecutul său. Dacă o bacterie detectează o anumită concentrație de molecule de lactoză în apa în care tocmai a înotat, este foarte adaptiv să poată compara aceste condiții cu o amintire a concentrației din mediul pe care tocmai l-a părăsit; comparația îi va determina mișcările ulterioare. Dacă o bacterie își amintește că tipurile de temperaturi ambientale pe care le experimentează în prezent au fost, în trecut, urmate în mod obișnuit de o creștere bruscă a calcificării, aceasta va fi foarte adaptată pentru a putea face o ajustare anticipată a metabolismului pentru a se adapta acestor circumstanțe viitoare.

Cele mai multe dintre aceste amintiri, precum cele referitoare la concentrațiile de nutrienți, sunt pe termen scurt, unele chiar de trei sau patru secunde, deși reținerea informațiilor poate dura mai mult de un minut. Acestea sunt amintiri pe termen scurt, dar sunt, literalmente, amintiri și foarte adaptative. Altele, precum cele care modifică expresia genetică, sunt pe termen lung. În 2017, Roland Mathis și Martin Ackermann de la Institutul Federal Elvețian de Știință și Tehnologie Acvatică din Zurich au raportat existența unui sistem de memorie nongenetică de durată mai lungă în bacteria *Caulobacter crescentus*. *C. crescentus* este o bacterie sesilă care se divide asimetric, celula „mamă” sesilă, cu pedunculi, rămânând pe locul său, iar celula „fiică” mobilă plecând în altă locație.

Mathis și Ackermann au supus celulele la mai multe valuri de o soluție concentrată de sare aversivă, înainte de diviziunea celulară. Apoi, după diviziune - și după relocarea celulelor fiice - au reintrodus sărurile toxice. Rata de mortalitate celulară rezultată în populația de celule mamă a fost

semnificativ mai scăzută decât în rândul celulelor fiice, iar acest efect profilactic a durat până la două ore. În mod clar, bacteriile care au rămas în același mediu au păstrat o înregistrare mentală a evenimentelor de acolo și s-au arătat pregătite să facă schimbări metabolice adecvate pentru a face față sărurilor toxice în cazul în care acestea se repetă. Celulele fiice mobile, cele care au plecat într-o locație nouă și probabil sigură, nu au avut nevoie să păstreze memoria evenimentelor anterioare și nu au reușit să facă aceleași tipuri de ajustări metabolice. Prin urmare, acestea au sucombat cu o rată mai mare, în ciuda faptului că aveau o compoziție genetică identică.

Și mai interesant, Mathis și Ackermann au raportat că bacteriile pot forma, de asemenea, ceea ce se poate numi doar o „memorie colectivă” cu o scară de timp destul de lungă. Atunci când au supus colonii mari de *C. crescentus* la valuri similare de soluții sărate concentrate, au găsit dovezi pentru o memorie a evenimentelor până la două ore mai târziu. Efectul pare să fie o caracteristică a mecanismului *quorum sensing* discutat în capitolul 2, în care toate celulele răspund la schimbarea bruscă a caracterului metabolic al mediului și reduc diviziunea celulară într-un mod sincronizat. Această schimbare a strategiei de reproducere modifică funcțiile metabolice în moduri care cresc ratele de supraviețuire, iar ratele de mortalitate celulară au rămas relativ scăzute în colectivități timp de două ore. Cu toate că două ore nu reprezintă o perioadă impresionantă dacă ne uităm la mamifere sau chiar la cefalopode, este destul de remarcabilă la o specie unicelulară, în special la una a cărei durată de viață se încadrează în acest interval general de timp.

Memoria imunologică.

Dacă sunt invadate de un virus sau de o plasmidă (un lanț de acid nucleic), bacteriile și alte procariote pot, folosind o proteină cunoscută sub numele de Cas9, să taie o parte din materialul genetic al intrusului și să o insereze în propriul ADN. Acest proces le permite să stabilească o reprezentare pe termen lung a identității intrusului și să fie mai bine pregătite pentru a face față oricărui atac viitor. Deoarece procedura modifică genele bacteriei, aceasta creează, de fapt, o memorie permanentă, care este transmisă în timpul diviziunii celulare. Această procedură este cea care stă la baza CRISPR, tehnica revoluționară de editare a genelor care permite cercetătorilor să taie secvențe specifice de ADN, să le mute în altă parte de-a lungul șirului de ADN și să introducă secvențe noi.[66]

CRISPR are potențialul de a vindeca tulburările genetice prin modificarea ADN-ului pentru a elimina genele mutante sau a introduce segmentele lipsă și este explorată de Jennifer Doudna, biochimist molecular la Universitatea Berkeley din California, ca tehnică potențială pentru stabilirea imunității virale la om. Doudna și colaboratoarea sa Emmanuelle Charpentier au realizat cea mai mare parte a lucrărilor inițiale privind CRISPR, dar mecanismul a fost, în mod prozaic, „inventat” de procariote.[67]

De asemenea, biologii informaticieni o folosesc într-un mod inedit pentru a transforma bacteriile în mici unități flash. Geneticianul George Church și colegii săi de la Harvard au reușit să introducă șiruri arbitrare de ADN într-o bacterie, unde acestea nu numai că devin parte a codului genetic al organismului, dar sunt transmise urmașilor săi în timpul diviziunii celulare. Șmecheria constă în utilizarea unor secvențe de ADN care seamănă cu cele ale unor viruși sau plasmide existente și în păcălirea bacteriei pentru a le încorpora. Până în prezent, Church și colegii săi au reușit să introducă până la 100 de octeți într-o singură bacterie *E. coli* și au folosit o colonie de bacterii pentru a înregistra un întreg film prin codificarea unei mici părți a filmului în fiecare celulă. Rămâne de văzut dacă acestea sunt ceva mai mult decât trucurile inteligente de salon ale unui om de știință, dar, în orice caz, ele dezvăluie o parte din capacitatea remarcabilă a *E. coli* de a codifica și

66 CRISPR vine de la Clustered Regularly Interspaced Short Palindromic Repeats (Repetiții palindromice scurte regulat intercalate). Vom reveni asupra acestei tehnici mai târziu, dar cititorii interesați pot consulta Wikipedia, care are o intrare lungă și detaliată despre această tehnică.

67 Philippe Horvath și colegii săi din Franța au fost primii care au demonstrat că o bacterie, S. thermophilic, poate dobândi rezistență împotriva unui invadator prin integrarea în genomul său a unui fragment de genom al virusului infecțios. A se vedea Barrangou et al. (2007) pentru detalii.

reține informații (a se vedea Shipman, Nivala, Macklis și Church, 2016.2017 pentru tehnicile utilizate).

O mare parte din această recentă recrudescență a interesului pentru aceste aspecte ale învățării și memoriei la bacterii este determinată nu atât de fascinația pentru bazele moleculare ale percepției sau ale învățării și memoriei la organismele unicelulare, ci de preocupări mai practice legate de rapiditatea și eficiența cu care unele dintre aceste bacterii extrem de virulente care provoacă atât de multă suferință umană dezvoltă rezistență la antibiotice. Este clar că înțelegerea acestor mecanisme ar putea oferi noi modalități de contracarare a impactului lor. Și, desigur, într-un fel de seducție captivantă, aceste noi investigații științifice asupra proceselor moleculare ale celor mai elementare funcții organice ne oferă o nouă platformă pentru a analiza ceea ce sunt, de fapt, probleme psihofilozofice.

Ar mai fi multe de spus pe această temă, dar cred că punctul de vedere a fost atins. Mecanismele care stau la baza acestor funcții de memorie sunt complexe. Ele implică schimbări rapide atât la nivelul receptorilor, cât și al moleculelor din celulă, care mențin o evidență curentă a condițiilor actuale, le coordonează cu reprezentările circumstanțelor întâlnite anterior și generează semnale către structurile din celulă care controlează mișcarea. Mecanismele precise implicate sunt, desigur, de interes mai ales pentru biologii moleculari, dar oricine este serios interesat poate consulta prezentarea din 2016 a lui Ganhi Lan și Yuai Tu a unora dintre mecanismele despre care se crede în prezent că funcționează aici, în special cele pe care *E. coli* le utilizează în sistemul său chemotaxic pentru a forma memorii de lucru. Punctul critic pentru abordarea **CBC** este că aceste sisteme sofisticate de învățare și memorie sunt prezente în toate speciile primordiale și fac parte din sensibilitatea acestor organisme unicelulare.

Luarea deciziilor.

Încă de la studiile inovatoare ale lui Jacques Monod privind modul în care *E. coli* trece de la o stare metabolică la alta în funcție de anumite zaharuri prezente în mediul său, biologii celulari au considerat bacteriile drept organisme care iau decizii. În cercetările inițiale efectuate în anii 1940, Monod a schimbat concentrațiile de zaharuri din mediul *E. coli* între glucoză (un zahăr preferat) și lactoză (nutritiv, dar nepreferat). El a constatat pauze în creșterea coloniei care corespundeau cu punctele în care toată glucoza era consumată și rămânea doar lactoza. Pauza s-a dovedit a fi cauzată de faptul că, în timp ce glucoza era abundentă, *E. coli* dezactivase gena care îi permitea să metabolizeze lactoza și avea nevoie de timp pentru a face ajustări în ADN-ul său care să îi permită să metabolizeze din nou eficient lactoza. Această descoperire a dat naștere unei adevărate industrii casnice care a explorat bazele moleculare și genetice ale acestor operațiuni. Monod a primit ulterior Premiul Nobel pentru Fiziologie și Medicină pentru cercetările sale privind mecanismele care controlează expresia genelor.

Întrebarea care a persistat (și care persistă și astăzi) este care este exact baza moleculară pe care o folosesc celulele pentru a „decide” (și da, acest cuvânt este cel folosit de cercetătorii din domeniu) să activeze sau să dezactiveze o genă. Într-o prezentare generală recentă a domeniului, biofizicianul Jane Kondev de la Universitatea Brandeis a examinat diferitele modele teoretice ale geneticii care stă la baza acestor funcții și a evaluat cât de bine a explicat fiecare datele privind funcțiile de selectare a dietei. Într-o secțiune inscripționată cu subtitlul amuzant „Are *E. coli* liber arbitru?”, a meditat:

Rezultatul este o populație de celule care, deși identice genetic [cu o alta], se pot comporta foarte diferit. Este ca și cum fiecare celulă E. coli este liberă să aleagă între două diete, glucoză sau lactoză, fără a fi constrânsă de genele sale sau de mediu. Experimentele... dezvăluie natura moleculară a acestui „liber arbitru” și îl trasează până la expresia stocastică a genelor E. coli.

Grupul de cercetare al lui Kondev se concentrează pe modelarea matematică, iar problemele care îi preocupă sunt elementele stohastice care permit prezența mai multor opțiuni, oricare dintre acestea putând fi luată de o bacterie, și ce distribuții decizionale apar atunci când sistemul face alegeri „libere” sau altfel neconstrânse. În abordarea lor a acestei probleme, circumstanțele care sunt prezente atunci când bacteriile aleg pot fi artificiale sau organice. Nu contează, deoarece se caută mecanisme generale, care funcționează atunci când mediul se schimbă, indiferent de factorii de schimbare.

Bănuiesc că Kondev nu ar dori să intre într-o dispută metafizică pe această temă cu filosofi precum Daniel Dennett, cu neuroștiința lui Sam Harris sau cu psihologi precum un alt prieten de-al meu care a murit prea devreme, regretatul Daniel Wegner[68], dar punctul său de vedere (cu care eu, evident, sunt de acord) este că posibilitatea de a alege între două sau mai multe moduri alternative de acțiune se încadrează bine în sensul psihologiei populare al *liberului arbitru*. Din punctul de vedere computațional al lui Kondev, bacteriile au liber arbitru.

Comunicarea.

După cum s-a menționat în capitolul 2, bacteriile sunt comunicatori eficienți atât în cadrul speciilor, cât și între specii. Nevoia de a interacționa ușor cu ceilalți apare deoarece bacteriile sunt, de asemenea, comunitare și formează cu ușurință colective cunoscute sub numele de covoare microbiene sau biofilme pe suprafețe precum dinții și coca navelor. Pe măsură ce colectivul începe să se formeze, bacteriile comunică între ele folosind mecanismul de detectare a cvorumului discutat în capitolul 2, în care moleculele cu funcții de semnalizare sunt secretate și modulează creșterea coloniei.[69]

Pe măsură ce colectivul crește, în cele din urmă apar probleme și competiția intercelulară devine un factor. Celulele de la periferie sunt cele mai apropiate de nutrienți, dar, deoarece sunt expuse mediului extern cu gama sa de prădători, cum ar fi antibioticele și moleculele toxice, sunt mai vulnerabile decât cele din interiorul colectivului. Cele din centru, deși se bucură de protecția împotriva prădătorilor pe care le-o oferă locația lor, sunt dezavantajate de nivelurile mai scăzute de nutrienți. Dacă numărul celulelor exterioare crește prea repede, membrii cei mai din interior ai colectivului vor muri de foame, la propriu. În mod clar, la un moment dat, comunitatea va ajunge la un punct de criză și va trebui să se facă ceva, altfel nucleul coloniei va muri, lăsând restul populației în pericol din cauza prădătorilor din toate părțile.

Ceea ce se întâmplă este o minune a evoluției și o fereastră către viața bacteriilor și poate primele indicii ale valorii adaptative a cooperării și altruismului. Atunci când celulele din centru detectează că nivelul nutrienților a scăzut prea mult, atât de mult încât viața lor este în pericol, ele produc un metabolit care semnalează necesitatea unei scăderi a ratei de creștere celulară. Pe scurt, ele le spun bacteriilor periferice să oprească diviziunea, să încetinească procesele metabolice și să lase o parte din hrană să se deplaseze spre interior. Celulele cele mai exterioare cooperează. Ele își ajustează ratele metabolice, consumă mai puține substanțe nutritive și reduc rata de diviziune celulară. În scurt timp, nutrienții se îndreaptă spre interiorul celulelor centrale. Când acestea își revin din seceta calorică, își modifică semnalele biochimice, iar celulele periferice încep din nou să înghită alimente și să se dividă cu abandon. Acest ciclu de faze continuă atâta timp cât colectivul supraviețuiește.

---

68 Cartea lui Wegner, *The Illusion of Conscious Will* (Iluzia voinței conștiente), este o analiză a noțiunii de voință care susține, în mod convingător în opinia mea, că nu avem nimic asemănător cu ceea ce omul obișnuit de pe stradă consideră a fi „liberul arbitru”. Ceea ce avem este o gamă mai banală, dar mai ușor de înțeles, de lucruri pe care putem alege să le facem sau nu, în anumite circumstanțe. Pe scurt, avem capacitatea de a alege, dar aceasta este departe de a fi „liberă” și, în cele din urmă, noțiunea de „voință” este cel mai bine tratată metaforic. Problema, desigur, este departe de a fi rezolvată (a se vedea Baggini, 2005).

69 În treacăt fie spus, merită menționat faptul că senzorul de cvorum este utilizat și de specii mai complexe, inclusiv de insectele stup, cum ar fi albinele și termitele, pentru a lua decizii cu privire la momentul în care un cuib și-a depășit dimensiunea optimă și ar trebui divizat.

În mod curios, atunci când colectivul creşte, comunicarea între bacterii se realizează prin mijloace electrice. Bacteriile pot simți când mărimea biomasei (care poate fi formată din mai multe tipuri diferite de bacterii) a devenit prea mare pentru o comunicare biologică simplă. Indivizii dinspre centru eliberează acum ioni de potasiu prin porii din membrana celulară. Aceste particule încărcate stimulează bacteriile vecine (de toate tipurile, deoarece acest mecanism are impact asupra organismelor unicelulare indiferent de specie) să le urmeze exemplul. Semnalul electric încărcat pozitiv se propagă rapid ca o undă în întregul grup de celule și toate bacteriile ajustează în mod corespunzător rata de diviziune celulară și încetinesc creşterea coloniei. Odată ce nivelurile de nutrienți s-au restabilit, eliberarea de ioni încetează și diviziunea celulară începe din nou (pentru detalii detaliate privind acest proces, a se vedea Beagle și Lockless. 2015).

În plus, după cum au descoperit Gürol Süel și colegii săi de la Universitatea din California, San Diego, lucrând cu bacteria *Bacillus subtilis*, această formă de comunicare permite, de asemenea, ca două colonii separate de bacterii să își coordoneze reciproc creşterea pe baza nivelurilor generale de nutrienți din mediul lor comun (a se vedea Prindle et al.. 2015 pentru articolul original). Cele două colonii folosesc ceea ce cercetătorii au numit o versiune a „împărțirii timpului". Colectivul care se confruntă cu niveluri diminuate de nutrienți trimite un semnal coloniei îndepărtate, folosind același sistem de semnalizare cu ioni de potasiu, solicitându-i să își ajusteze rata metabolică și frecvența diviziunii celulare. Răspunsul este, aproape invariabil, unul acomodativ. Este ca și cum colectivul care se bucură în prezent de abundența nutritivă apreciază virtuțile altruismului - cu înțelegerea implicită că ambele părți recunosc contractul și sunt de acord să îl respecte. Și chiar mai remarcabil, efectul a fost observat recent între diferite specii de bacterii (a se vedea raportul din Humphries et al.. 2017) - un exemplu clasic de altruism neamical.

Motivație și afecte.

Speciile unicelulare prezintă toate semnele unor organisme cu mecanisme motivaționale puternice. Și, așa cum se întâmplă cu toate comportamentele motivate, acestea se bazează pe afect, pe sentimente, pe emoții. Deoarece viața organismelor unicelulare se concentrează pe cele mai elementare funcții, reproducerea, hrănirea, excreția și evitarea moleculelor virulente, în esență, fiecare aspect al acțiunilor lor este determinat de nevoi biologice de bază.

Am abordat mai devreme distincția lui Antonio Damasio între experiența subiectivă a ceea ce el a numit „emoție" și cea a „sentimentului". În nomenclatura sa, prima este o trăire mai primitivă a unui eveniment afectiv, în timp ce a doua are un element interpretativ.[70] Această distincție este recomandabilă. În numeroasele exemple de sensibilitate a procariotelor pe care le-am analizat în acest capitol, probabil că este mai bine să le vedem ca instanțieri ale unor stări interne care exprimă ceva asemănător unei „emoții" damasiene - cu înțelegerea faptului că aceste stări interne sunt cele însoțite de conștiință, de sensibilitate adevărată (un punct de vedere cu care Damasio nu este de acord, este corect să observăm). Fiecare stare de nevoie motivațională este însoțită de reprezentări interne, subiective, fenomenale. Nu este nevoie să repetăm afirmațiile anterioare; rețineți doar că aproape fiecare funcție pe care am discutat-o, de la percepție la luarea deciziilor și comunicare, se bazează pe faptul că organismele unicelulare sunt conștiente și receptive la valențele obiectelor pe care le întâlnesc, la experiențele lor afective cu acestea și la alegerile pe care le fac pentru a se adapta la ele.

Bacteriile dorm?

70 În treacăt fie spus, rețineți că etichetele lui Damasio nu sunt în conformitate cu modul în care mulți le folosesc. În stilul „psihologiei populare" pe care îl folosim în acest volum, emoțiile sunt considerate, în general, instanțieri mai sofisticate, în timp ce sentimentele sunt considerate mai elementare și mai puțin interpretate. Nu sunt singurul care consideră terminologia lui Damasio incomodă; alții, inclusiv cercetătorul cognitiv Stevan Hamad și filosoful Galen Strawson, și-au exprimat opinii similare.

O întrebare rezonabilă, având în vedere că una dintre caracteristicile vieții este să ia o pauză, permițând organismului să se recupereze și să restabilească funcțiile metabolice normale, să repare leziunile celulare. Cele mai bune dovezi de până acum sugerează că răspunsul este probabil nu, cu siguranță nu în modul în care suntem obișnuiți să vorbim despre somn ca despre o stare cu caracteristici neuronale specifice, marcate de schimbări în electroencefalograme (EEG) și cu faze distincte, cum ar fi somnul cu mișcări oculare rapide (REM), pe care îl vedem la animalele cu sistem nervos. Ne-am aștepta cu greu să vedem aceste mecanisme în funcțiune în organismele preneurale.

Cu toate acestea, unele specii de bacterii sunt sensibile la variațiile nivelului de lumină ambientală și prezintă ritmuri circadiene. Dintre procariote, grupul cel mai studiat este cel al *cianobacteriilor.* După cum au demonstrat Susan Golden și colegii săi de la Universitatea California din San Diego (pentru detalii, a se vedea Boyd et al.. 2016), cianobacteria *Synechococcus elongatus* modifică ratele de exprimare a genelor și procesele fiziologice ca răspuns la ciclurile de lumină și întuneric. Cianobacteriile sunt deosebit de interesante deoarece utilizează fotosinteza pentru a genera energie și se presupune că au fost implicate în evenimentul chimeric care a avut loc acum aproximativ 1,7 miliarde de ani, când au evoluat primii precursori ai regnului vegetal. Vom avea mai multe de spus despre acest eveniment singular atunci când vom analiza mai îndeaproape întrebarea dacă plantele pot fi considerate a fi sensibile.

Sentimentul colectiv.

După cum probabil ați observat, câteva dintre discuțiile de mai sus au implicat examinarea comportamentului unui colectiv de organisme unicelulare. Acesta a fost cazul învățării multigenerațională și al efectelor memoriei, al comunicării care utilizează atât mecanisme de detectare a *quorum*-ului, cât și mecanisme de eliberare a ionilor de potasiu, precum și al reacțiilor afective la atractanți și repelenți.

Aici se pune o problemă importantă, una care se pune în evidență în biologia evolutivă și, în virtutea ipotezelor de bază ale modelului **CBC**, una care va trebui rezolvată în cele din urmă. Cum ajung speciile pluricelulare să aibă experiențe integrate și unitare atunci când sunt, în esență, conglomerate de multe celule biologic distincte, fiecare dintre acestea, conform principiilor de bază ale modelului **CBC**, fiind sau fiind ea însăși conștientă? Rezultatele studiilor lui Suel și cele ale lui Tavazoie și ale colegilor discutate anterior sugerează că prima apariție a acestei schimbări epistemice de la organismul individual la un colectiv are loc atunci când bacteriile individuale se agregă pentru a forma biofilme, iar nevoile coloniei ajung să le înlocuiască pe cele ale indivizilor care o constituie. Vom avea mai multe de spus despre acest aspect în capitolul următor. Este o adevărată provocare și invită la speculații cu privire la ceea ce se întâmplă în lumea microorganismelor.

Rezumat

În literatura științifică de micro-biologie și biologie celulară, explorările acestor numeroase funcții perceptive, motorii, cognitive și afective sunt de obicei prezentate în limbajul tehnic, lipsit de pasiune, al oamenilor de știință obiectivi. Câteva ating subiecte precum conștiința sau o anumită formă de sensibilitate, iar câțiva autori au folosit termeni precum „decizie” și „liber arbitru”, dar în mare parte este ca în orice alt domeniu științific: context, teorie, metodă, rezultate, interpretare. Cu cât am cercetat mai mult literatura de specialitate, cu atât lucrurile mi s-au părut mai simple. Cele mai simple organisme experimentează, simt, detectează și interpretează evenimente; își modifică comportamentele; își amintesc experiențele trecute; și fac alegeri. Aceste funcții le oferă un sens al

ceea ce este endogen fiecăruia și le permite să distingă fiecare fragment de conținut epistemic de informațiile despre lumea exterioară, inclusiv alte organisme din mediul imediat. Pe scurt, bacteriile și alte specii unicelulare disting sinele de non-eu - o caracteristică a conștiinței frecvent menționată.

Nu sugerez că *E. coli* are discuții introspective cu ea însăși cu privire la planurile sale pentru restul după-amiezii, dar apăr ipoteza că are reprezentare de sine, este conștientă de stările metabolice interne, are cunoștințe despre natura biochimică, moleculară a mediului său, utilizează informații homeostatice pentru a-și ghida alegerile și are experiențe fenomenale primitive, dar foarte reale ale acestor procese.

În prelegerea sa pentru Nobel din 1983, eminentul genetician Barbara McClintock a trecut în revistă numeroasele funcții remarcabile ale speciilor unicelulare așa cum erau ele înțelese atunci, în special adaptările pe care le fac la complexitatea și fluctuațiile mediului lor. Ea a spus apoi într-un mod sugestiv că, din cauza numărului mare de adaptări pe care trebuie să le facă,

Dispozitivele de detectare și semnalele care inițiază aceste ajustări sunt dincolo de capacitatea noastră actuală de a le înțelege. Un obiectiv pentru viitor ar fi acela de a determina gradul de cunoaștere pe care celula îl are despre ea însăși și modul în care utilizează aceste cunoștințe într-un mod „atent" atunci când este provocată. (Subliniere adăugată.)

Luate împreună, aceste funcții pe care tocmai le-am trecut în revistă surprind ceea ce McClintock se simțea confortabil să numească cunoștințe și să afirme că utilizarea lor este un act de gândire. Microbiologii Sharad Ramanathan și James Broach au fost de acord. Aceștia au observat într-o lucrare din 2007 că o mare parte din ceea ce se întâmplă în speciile unicelulare este echivalentul funcțional al gândirii. De fapt, titlul acelei lucrări era "*Do Cells Think?*" (Gândesc celulele?). Ramanathan și Broach au subliniat că există două moduri în care celulele ar putea face față fluctuațiilor naturale ale moleculelor critice care abundă în lumea lor - și ambele implică adaptări pentru a anticipa schimbarea. O strategie ar fi să anticipeze schimbarea prin producerea de variații fenotipice multiple - o strategie pe care am întâlnit-o deja. Este cea pe care Gerald Edelman a descoperit-o funcționând în sistemul imunitar, unde se produce un număr mare de tipuri diferite de limfocite, pregătite practic pentru orice antigen care s-ar putea ivi. Cealaltă este capacitatea de a identifica schimbările pe măsură ce acestea apar și de a modifica comportamentul acționând prin deplasare, modificarea permeabilității membranei celulare și creșterea sau scăderea ratelor de diviziune celulară. După ce au analizat instanțierea acestor strategii în diferite specii unicelulare, Ramanathan și Broach au concluzionat:

Aplicarea cu succes a ambelor strategii necesită un nivel de memorie și de prelucrare a informațiilor care nu a fost asociat în mod normal cu celulele individuale, ceea ce sugerează că aceste organisme au, de fapt, capacitatea de a „gândi".

În mod curios, laboratorul lui Ramanathan de la Harvard își definește obiectivul principal ca fiind descoperirea mecanismelor care stau la baza luării deciziilor și alegerii. Acest paragraf este preluat direct de pe site-ul său:

Laboratorul nostru studiază modul în care celulele și organismele iau decizii. Pentru a ajunge la o decizie, organismele trebuie să măsoare mai multe semnale de mediu și să le interpreteze în mod corespunzător. Întrebările care ne interesează sunt modul în care celulele și organismele își interpretează mediul, modul în care interpretarea depinde de experiențele anterioare, precum și modularea spațială, temporală și statistică a semnalelor de mediu.

Eliminați cuvântul „celule" din acest paragraf și programul de cercetare al lui Ramanathan ar putea descrie cu ușurință laboratorul unui ecologist cognitiv care explorează funcționarea socială a cimpanzeilor sau programul de cercetare al unui psiholog comparativ care analizează modul în care corvidele rezolvă problemele folosind amintirile experiențelor anterioare pentru a se ghida.

Pe scurt, viața creează conștiința, și o face fără un sistem nervos. Acolo, la început, este o existență mentală foarte mică, dar determinabilă. O bacterie unicelulară sau un *arhean* are subiectivitate, o minte, conștiință și un simț al sinelui. Aceasta este cheia, axioma fundamentală a teoriei **CBC**. Chiar trebuie să fie așa, trebuie să existe o sensibilitate primitivă aici, la acest nivel de viață, altfel strămoșii noștri originari nu ar fi putut supraviețui, nu ar fi putut naviga în mediile lor pline de flux, schimbare și pericol. Dacă vrem să acoperim golul explicativ, trebuie să ne uităm aici. Acesta este momentul în care materia și-a dat seama cum să creeze mentalul. Toate celelalte forme de sensibilitate sunt variații pe o temă, această temă.

Coda: Anticiparea criticilor (încă o dată) și apărarea împotriva lor

În acest moment, eu, noi, trebuie să punem câteva întrebări, întrebări la care am făcut aluzie de mai multe ori, dar care, acum că am analizat dovezile existenței celulelor sensibile, trebuie abordate din nou: Toate cele de mai sus sunt doar antropomorfism dezlănțuit? Nu ar putea toate aceste comportamente, învățate și uitate, să fie simple acțiuni și reacții oarbe, codificate genetic? Sunt toate aceste talente unicelulare doar exemple de „competență” lipsită de „înțelegere” - toate procesările de jos în sus fără o licărire de subiectivitate, de sensibilitate? Totul este doar „a face” fără „a simți”? Ar putea un filozof cu o înclinație diferită de a mea, cum ar fi Dan Dennett, să privească această serie de bacterii inteligente și protozoare sesile care rezolvă probleme și să spună (așa cum a făcut el, parafrazând un schimb de e-mailuri): „*Ei bine, profesore Reber, încă nu mi-ați arătat nimic suficient de convingător pentru a mă face să cred că toate aceste comportamente pe care le-ați expus cu atâta dragoste în fața noastră sunt ceva mai mult decât simple procese de rutină, stimul-răspuns, care au fost programate în genele lor. Nu văd nimic aici care să semene nici pe departe cu o viață mentală autentică.*”

Am încercat să abordez această critică mai devreme, dar trebuie să o fac din nou, deoarece, atunci când este dezbrăcată, este o poziție care, susțin eu, nu câștigă nimic în putere explicativă și este plină de dificultăți logice, filosofice și biologice evolutive:

> În primul rând, adoptarea poziției lui Dennett, așa cum am observat, nu ajută cu adevărat; de fapt, vă pune în fața unei probleme mai profunde. Dacă doriți să argumentați că mintea și sensibilitatea nu au păşit în fața reflectoarelor existențiale la începuturile vieții, trebuie să explicați cum și de ce mecanisme adaptative atât de sofisticate ar fi putut evolua în absența acestor stări subiective. Ce rost are să avem ceea ce arată și funcționează exact ca un sistem senzorial dacă acesta nu furnizează informații subiective utile și utilizabile speciei care îl posedă? Ce cale darwiniană ar fi putut evolua pentru construirea unui sistem metabolic care poate ține evidența evenimentelor trecute și le poate păstra suficient de mult timp pentru a lua decizii măsurate cu privire la acțiunile și mișcările pe care trebuie să le întreprindă, dacă nu există o reprezentare internă a acestor evenimente? De ce ar trebui protoplasma să fie sensibilă la gradienții de nutrienți din apa de mare dacă nu există o instanțiere încărcată cu valențe a acestora? De ce să existe o mișcare controlată dacă nu există o conștientizare a locului și a momentului în care trebuie să se miște?
>
> Mai rău, fiecare dintre aceste comportamente și o serie de alte comportamente ar necesita, din punctul de vedere al lui Dennett[71], un mecanism separat, codificat genetic, pentru a le explica - unul pentru mișcare, unul pentru detectarea nutrienților, unul pentru temperatură, altul pentru toxine, încă unul pentru prădători și încă unul pentru detectarea

71 Mai devreme l-am criticat pe Chalmers. Aici mă iau de Dennett. Am făcut schimb de opinii cu Dan cu privire la aceste probleme. Îl cunosc de la acel institut de vară fundamental din 1981. Îl consider un prieten și un combatant demn. Eu cred că se înșeală aici. El aproape sigur crede că eu greșesc. Așa stau lucrurile - și așa ar trebui.

cvorumului şi acţiunea colectivă. Susţin că este extrem de puţin probabil ca acestea să fie procesele de selecţie naturală prin care au evoluat comportamentele complexe - în special pentru că acestea ar trebui să fie funcţionale în timp ce se deplasează într-un mediu în continuă schimbare. Schimbarea, fluxul, variaţia sunt elemente de bază ale lumii. „Competenţele" genetice programate nu se descurcă bine, dacă nu deloc, cu circumstanţele schimbătoare. Pe de altă parte, adoptând poziţia recomandată de perspectiva **CBC**, toate aceste procese, toate aceste funcţii, pot fi privite ca produse directe ale unui singur factor cauzal: *sensibilitatea*. Încercarea de a explica comportamentele speciilor unicelulare ca o serie de acţiuni separate, determinate genetic şi lipsite de minte este complicată, greoaie şi extrem de improbabilă.

În al doilea rând, presupunerea că aceste tipuri de sisteme fără minte, asemănătoare roboţilor, au evoluat independent de subiectivitate nu ajută, câtuşi de puţin. Nu oferă nicio putere explicativă suplimentară. Odată ce adoptăm această poziţie „lipsită de minte", ne întoarcem, înfipţi în ţepuşa *emergentismului*. Din moment ce ştim că oamenii au conştiinţă, minte, subiectivitate şi cunoaştere de sine, tot va trebui să elaborăm un set de circumstanţe în care a apărut pentru prima dată prototipul experienţei fenomenale. Şi această problemă, după cum am văzut deja, este o nucă foarte greu de spart. De fapt, o parte importantă a poveştii pe care încerc să o spun aici este că este o nucă nepotrivită pe care să o rozi în primul rând.

În al treilea rând, chiar dacă doriţi să aruncaţi teoria **CBC** în tomberonul epistemic şi să continuaţi să încercaţi să rezolvaţi problema extrem de refractară, veţi fi împotmoliţi în aceleaşi gropi de gudron logice în care se găsesc blocaţi cercetătorii actuali în filosofia minţii. Vă veţi învrăjbi să căutaţi soluţii complicate, întortocheate şi inflexibile. **CBC**, pe de altă parte, oferă o nucă mult mai uşor de spart. Desfaceţi mecanismele microbiologice care au dat naştere apariţiei iniţiale a sensibilităţii - care, sunt destul de sigur, vor fi strâns legate de cele care au dat naştere apariţiei iniţiale a vieţii. Poate că nu suntem nici pe departe pe punctul de a „rezolva" *problema dificilă*, dar abordarea **CBC** oferă o modalitate de a o „re-rezolva". În plus, ne putem alinia frumos cu Sir William de Occam, Conwy Lloyd Morgan şi toţi apărătorii unei parcimonii pragmatice. Şi, extinzând metaforele noastre cu vopsea umedă şi groapă de gudron lipicios până la punctul de ruptură, vom fi tăiat unul dintre multele noduri gordiene.

În al patrulea rând, vom obţine o realiniere a priorităţilor în neuroştiinţele cognitive. În loc să ne străduim din răsputeri să dezlegăm misterul modului în care creierul creează mintea, putem continua să facem ştiinţă fundamentală şi să lucrăm la probleme uşor de rezolvat, cum ar fi rolul ganglionilor bazali în învăţarea implicită sau relaţiile dintre densitatea celulelor lobului frontal şi conştiinţă, sau rolul centrelor cerebrale multiple la cefalopode, cortexul vizual al albinei *bumble*, detectarea cvorumului în comportamentul roiurilor, sau să continuăm să lucrăm la oricare dintre descoperirile recente pe care filozoful Ned Block şi coautorii le-au citat ca provenind din ceea ce ei au numit „ştiinţa conştiinţei":

... atenţia şi conştiinţa sunt disociabile, atenţia de diferite tipuri poate funcţiona în absenţa conştiinţei, există ... experienţă conştientă fără atenţie sau raportare ... o înţelegere a mecanismelor neuronale şi cognitive ale metacogniţiei, o înţelegere a propriei conştiinţe şi performanţe ... o înţelegere a activităţii cerebrale asociate cu diferite niveluri de conştiinţă, inclusiv dezvoltarea de metode de detectare a prezenţei acesteia în somnul profund, ... pacienţi anesteziaţi şi cu deficienţe neurologice, ... practici mai bune de diagnosticare a tulburărilor de conştiinţă şi ... conştiinţă păstrată la pacienţii în stare vegetativă.

Aceste tipuri de descoperiri - cele care au fost făcute în ultimele decenii în domeniul cognitiv şi al neuroştiinţelor - sunt, în mod interesant, printre cele pe care Chalmers le-a enumerat drept „probleme uşoare". Nimic din modelul **CBC** nu contracarează interesul pentru inteligenţa artificială, microbiologia şi biologia celulară sau neuroştiinţele cognitive. Dimpotrivă, le stimulează, încurajează continuarea cercetării, deoarece aceasta poate fi

efectuată fără a fi afectată de preocupările persistente cu privire la modul de abordare a *problemei dificile*. Din punctul de vedere al **CBC**, **problema dificilă** este *dizolvată*; ea dispare, iar noi putem fi mulțumiți să lucrăm la toate cele ușoare.

În al cincilea rând, *misterianismul*, care și-a ridicat capul invizibil și greșit în ultima vreme, este împins în afara scenei - acolo unde îi este locul. **CBC** recomandă o întoarcere la un fizicalism de bază, la un reducționism pragmatic și la o îmbrățișare neclintită a empirismului. Acest ultim punct este important. S-a scris atât de mult despre insolvabilitatea problemei dificile, despre acele lacune explicative, despre limitările perspectivei la persoana întâi, despre inutilitatea de a desluși vreodată misterele minții, și lipsa de speranță de a descoperi modul în care simpla materie creează mentalul, încât începe să pară că această căutare nebunească pentru a înțelege ce înseamnă să ai conștiință este condusă mai puțin de știință și filozofie solidă și mai mult de un fior al misterului care, la fel ca și contemplarea unui Dumnezeu omniscient, a unui univers infinit sau a unei esențe inefabile, ne scapă printre degete, ca argintul viu, eludându-ne înțelegerea.

În această lume a noului mister, decalajul explicativ devine cumva real și de netrecut; imposibilitatea materiei de a crea mintea este considerată o fatalitate epistemică, acceptată ca evanghelie, îmbrățișată ca adevăr fără dovezi. **Problema dificilă** este transformată în **Problema imposibil de rezolvat** sau în **Cea mai dificilă problemă** dintre toate. Este aproape ca și cum filosofii care evită materialismul pentru una sau alta dintre pozițiile misteriene tânjesc după o spiritualitate pe care au abandonat-o în favoarea unei științe agnostice care nu îi satisface.

Abordarea **CBC** ne salvează de acest tip de noțiuni misteriene. Ea deschide o altă cale empirică pe care să rătăcim, dar este una care are un obiectiv bine definit la final: descoperirea proceselor biochimice, moleculare care au apărut pentru prima dată în cele mai simple forme de viață și care au produs apariția sensibilității. După aceea, va fi doar mai multă biologie evolutivă, ceva ce facem destul de bine.

5. Implicații ale bazei celulare a poziției conștiinței

Acest capitol ne va duce la sfârșitul călătoriei noastre. Prefața a început cu o notă personală, așa că haideți să avem încă una.

Când am început acest proiect, care datează de la mijlocul anilor 1990, eram destul de sigur că principiile de bază care stau la baza abordării *Bazei Celulare a Conștiinței* (**CBC**) erau corecte, că primele forme de viață aveau experiențe subiective, *qualia*, o conștiință ontologic sigură. Dar acesta era un articol de credință. Nu cercetasem literatura de biologie celulară; chiar nu știam ce se află acolo și eram pregătit să abandonez proiectul dacă datele nu se dezvoltau într-o manieră favorabilă. Ca să fiu sincer, pe atunci nu știam care este diferența dintre procariote și eucariote, nu auzisem niciodată de *Stentor roeseli* și cu siguranță nu știam să scriu pe litere partea „E” din *E. coli* (încă nu știu).

Ceea ce am descoperit atunci când mi-am revenit a fost, așa cum tocmai am analizat, uimitor și interesant. Dovezile existenței sensibilității unicelulare sunt copleșitoare. Nu sunt singur în această afirmație; o serie de microbiologi, neuroștiințifici, biologi evoluționiști și filosofi proeminenți, precum George Gaylord Simpson, Lynn Margulis, Barbara McClintock, Oliver Sacks, Maxine Sheets-Johnstone, Sharad Ramanathan, Frantisek Baluska, Dennis Bray - și cu siguranță alții ale căror contribuții la această problemă le-am ratat cu siguranță în peregrinările mele prin literatura cu adevărat enormă - și-au exprimat cu toții, într-un fel sau altul, convingerea că viața unicelulară este o viață sensibilă. La fiecare pas, fie că examinăm mecanismele senzoriale, abilitățile de interpretare a percepției, învățarea, memoria sau comunicarea, această lume este extrem de adaptivă și reușită. Esența, însăși ființa microorganismului, este una care înoată în sensibilitate, plină de experiență subiectivă, bogată în „cum este să fii” una dintre aceste creaturi surprinzătoare - chiar dacă recunoaștem punctul esențial al lui Nagel că nu putem avea niciodată o reprezentare la prima persoană a acestei stări.

De asemenea, am constatat în literatura tehnică o deconectare comună între biologii moleculari și celulari, pe de o parte, și filosofii minții, pe de altă parte. A fost o diviziune asemănătoare cu cea pe care am văzut-o în capitolul 1, când am analizat contribuțiile specialiștilor în robotică și ale cercetătorilor din domeniul inteligenței artificiale (AI) și relația lor cu filosofii. Oamenii de știință din ambele domenii, în cea mai mare parte, nu sunt foarte preocupați de implicațiile pe care descoperirile lor le-ar putea avea asupra problemelor care îi interesează pe filosofi. Oamenii concentrați în laborator, orientați empiric, care se bazează pe date, își văd în principal de treabă. Gurzii inteligenței artificiale scriu software, construiesc roboți, dezvoltă rețele neuronale multistrat bazate pe retropropulsie, sisteme de învățare profundă cu autocorectare, fără să se gândească dacă una dintre creațiile lor ar putea deveni sau este deja conștientă. În laboratoarele de biologie celulară se deslușesc funcțiile enzimelor, se explorează mecanismele de exprimare a genelor, trucurile de eliberare a ionilor care permit bacteriilor să comunice cu colonii îndepărtate de conspecifice și se utilizează tehnici de secționare a genelor procariotelor în efortul de a dezvolta o imunitate virală indusă. Și fac toate acestea fără să menționeze dacă rezultatele lor se referă la întrebarea cum ar putea materia să creeze mentalul.

Însă a apărut un aspect comun al vieții. Puținii oameni de știință care și-au îndreptat atenția către problema originii și naturii conștiinței, care și-au îndreptat atenția către examinarea problemelor ontologice mai ample ale minții și materiei, au făcut acest lucru mai târziu în cariera lor. Această cohortă include cercetători precum Francis Crick, Roger Penrose, Gerald Edelman, Antonio Damasio, Christof Koch și Giulio Tononi. Și, bineînțeles, subsemnatul urmează, de asemenea, acest traseu de viață.

Există cu siguranță multe motive pentru acest model. Unul care pare destul de evident este că, la începutul carierei noastre, cei mai mulți dintre noi am fost prinși în cercetări empirice foarte bine concentrate, încercând să răspundem la întrebări care se pretează mai bine la teorie, la examinarea experimentală și la o mulțime de date decât la argumente speculative sau la logica decupajului. De asemenea, bănuiesc că, pe măsură ce deceniile treceau, ne-am fi simțit mai puțin constrânși să facem ceea ce Thomas Kuhn numea „știință normală” și mai liberi să explorăm

probleme mai ample. Știu că acești factori sunt importanți în viața mea. A trecut aproape un sfert de secol de când am început să mă gândesc la o teorie a conștiinței bazată pe o singură celulă, dar am fost prea ocupat în acești ani intermediari pentru a face mai mult decât câteva propuneri „proclamatorii".

De asemenea, mă întreb dacă acei ani în care am făcut cercetări solide în domenii dificile și am obținut cel puțin o oarecare recunoaștere din partea colegilor nu ne-au dat un plus de încredere, de genul celui de care avem nevoie atunci când decidem să abordăm probleme dificile, în special cele precum originile minții, care poartă cu ele o notă de uimire și măreție. Deși nu dispun de date biografice, din peregrinările mele prin literatura de specialitate am avut sentimentul că relativ puțini dintre acești oameni de știință care s-au aventurat în studiul conștiinței și al minții au avut o pregătire formală în filosofie, în special în filosofia minții. În cazul meu, așa cum am menționat în prefață, experiența mea s-a limitat la un curs și la o vară de șase săptămâni de aprofundare în psihologia și filosofia minții. Totuși, așa cum am exprimat mai devreme, cred că această lipsă de expunere formală la literatură poate fi un avantaj. Când deschizi pentru prima dată ușa scârțâitoare a acestui seif de dezbateri sofisticate, nu ești la fel de captivat de toate disputele în curs ca cineva care a trecut prin rigorile unei formări avansate în domeniu. Faptul că nu știi prea multe despre filosofia minții ar putea oferi cuiva cu viziune un avantaj pe care nu l-ai crede. Puteți pur și simplu să adoptați, să zicem, o perspectivă materialistă sau să apărați un reducționism pragmatic fără să vă faceți griji cu privire la care dintre cele câteva versiuni ale acestor doctrine sunteți în favoarea sau în defavoarea dumneavoastră sau care sunt contra-argumentele obscure din literatura de specialitate.

Dar, nu contează. Indiferent dacă cineva este instruit în formele nuanțate ale discursului filosofic sau nu, în cele din urmă, desigur, este ca studiul conștiinței în sine. Suntem cu toții blocați cu perspectiva noastră de la prima persoană - știu doar ce m-a motivat să-mi îndrept atenția către problema enervantă a originii conștiinței și să dezvolt modelul **CBC**. Au fost, desigur, multe, dar în centrul lor a fost una pur și simplu personală: Cred că perspectiva **CBC** are implicații științifice și filosofice care merită luate în considerare. În esența sa, este o teorie științifică cu ipoteze complet testabile despre cum, de ce și când a apărut sensibilitatea, dar a fost elaborată într-un cadru filosofic. În timp ce **CBC** oferă biologilor celulari o modalitate inedită de a descoperi mecanismele care au dat naștere subiectivității, ea prezintă, de asemenea, o viziune filosofică care ar putea scoate filosofii minții din încurcătura creată de fascinația lor pentru **problema dificilă**.

Despre structura ierarhică a evoluției

În mai multe rânduri, am spus că evoluția nu renunță la adaptările utile. Aceasta se dovedește a fi o caracteristică esențială a evoluției și, atunci când este combinată cu o serie de principii conexe, conduce la o serie de perspective generale asupra trăsăturilor biologice și a modului în care acestea sunt instanțiate în speciile din întreaga gamă a vieții. După cum vom vedea, aceste principii ne permit, de asemenea, să apreciem cât de convingătoare este natura ierarhică a evoluției și modul în care aceasta reflectă o formă piramidală generală, cele mai vechi trăsături aflându-se la baza largă și fermă, iar cele mai recent evoluate spre vârf.

Următoarea prezentare este o versiune simplificată a unui model care a fost prezentat într-o varietate de lucrări publicate. Principiile sunt derivate din axiomele de bază ale teoriei evoluționiste și sunt înglobate în contextul sintezei moderne. După cum am subliniat în prefață, această abordare este un exemplu de utilizare a propozițiilor de bază din cadrul biologiei evolutive pentru a construi o teorie. Acesta nu este un exercițiu de psihologie evoluționistă și diferă dramatic de tipul de modelare văzut acolo. Oricine este interesat de modul în care s-a ajuns la următoarele principii poate găsi derivările formale în Reber (1993), Schank și Wimsatt (1988), și Wimsatt (1986).

- **Succes**. Odată ce apar forme de succes, acestea se stabilesc în acele specii care le prezintă. Toată evoluția este o mărturie a acestui principiu.

- **Conservarea**. Procesele evolutive sunt conservatoare. Odată ce formele de succes sunt stabilite, acestea tind să devină fixe și să servească drept fundamente pentru formele apărute ulterior.
- **Stabilitatea**. Funcțiile și formele adaptative apărute mai devreme tind spre stabilitate, prezentând mai puține variații de succes decât formele apărute mai târziu. Acest principiu este reflectat în sistemele de clasificare cladiste din biologia evolutivă, care se bazează pe existența unei specii ancestrale din care descind membrii cladei.
- **Comunitate**. Formele și funcțiile care au apărut mai devreme vor fi prezente la toate speciile. Cele care au evoluat mai târziu tind să fie specifice speciei care le prezintă. Acest principiu stă, de asemenea, la baza unei mari părți a etologiei și psihologiei comparative și, desigur, a abordării **CBC**.

Aceste principii generale ar trebui privite ca euristici ale cercetării și dezvoltării de teorii în cadrul biologiei evolutive. Cu alte cuvinte, ele nu sunt legi în sens strict și nu ar trebui considerate inviolabile. Ele sunt mai degrabă principii operaționale și, în general, valoarea lor derivă din capacitatea lor de a ghida cercetarea și teoria. Și fac acest lucru aici. Deoarece funcțiile senzitive ale organismelor unicelulare pe care le susține modelul **CBC** se numără printre primele caracteristici adaptative de succes din evoluție, comportamentul lor va reflecta toate aceste principii. Prin urmare, se presupune că *sensibilitatea* unui procariot ar fi:

- la fel de reușită ca orice altă trăsătură din istoria evoluției,
- conservată și regăsită în aproape toate speciile[72] care au evoluat ulterior,
- stabilă și servind drept bază pentru versiunile de sensibilitate și conștiință care au evoluat ulterior și,
- reflectând elemente comune prin faptul că trăsăturile sale de funcții senzoriale, perceptive, achiziționale, memoriale și motorii - împreună cu o reprezentare a sinelui, se regăsesc în instanțieri similare în întregul regn animal.

Pe scurt, structura evolutivă a conștiinței este ierarhică și piramidală. Funcțiile sale sunt fundamentale și adaptative. Versiunile care au evoluat ulterior au dezvoltat funcții și caracteristici noi. Acestea vor fi împărtășite de un număr mare de specii, deși fiecare versiune nouă, mai sofisticată din punct de vedere cognitiv, va fi instanțiată în mai puține clade distincte. De aici și comparația cu piramida. Conștiința umană se află la zenit, doar o singură specie prezentând toate caracteristicile sale, de la cele mai elementare, pe care le împărtășim cu procariotele, până la funcțiile cognitive și afective sofisticate care sunt exclusiv umane.

Rețineți, de asemenea, că trăsăturile de la bază trebuie să fie robuste și relativ invariabile, dat fiind că fiecare adaptare nouă bazată pe acestea nu poate avea succes dacă cea mai veche are chiar și mutații maladaptative ocazionale. Coloana vertebrală este un bun exemplu în acest sens. Odată ce vertebra a evoluat, a existat o morfologie clară și comună a scheletului în toate cele peste șaizeci de mii de specii de *Vertebrata*, inclusiv cele distincte precum amfibienii, reptilele, peștii, păsările și mamiferele. Modificările minore ale dimensiunilor și ale conturului osos sunt frecvente, dar forma de bază nu este alterată. Stabilitatea este menținută deoarece, practic, orice mutație în genele care controlează forma și/sau funcțiile coloanei vertebrale este de obicei fatală. *Hagfish*, o creatură foarte neplăcută, asemănătoare anghilei, cunoscută și sub numele de „anghilă de noroi”, este singurul caz cunoscut de animal cu craniu, dar fără vertebre.[73]

Prin urmare, caracteristicile străvechi, dar foarte adaptative, își mențin caracteristicile esențiale de-a lungul timpului geologic și între specii. De fapt, am întâlnit deja acest model în

72 "Practic" trebuie introdus din cauza plantelor. Mai târziu, în acest capitol, vom arunca o privire mai atentă și vom vedea că problema este atât complexă, cât și, până acum, nerezolvată.

73 A se vedea Tanvier (2010) pentru analize genetice ale *hagfish* și ale lampreilor, care arată în mod concludent că *hagfish* și-a pierdut într-adevăr coloana vertebrală. Este demn de remarcat faptul că această renunțare la o structură de succes este atât de rară încât Janvier s-a referit la ea ca la un exemplu de „degenerescență spectaculoasă”.

timpul discuțiilor noastre despre vedere, când am observat că *rodopsina*, unul dintre pigmenții fotosensibili din ochiul nostru, funcționează, de asemenea, ca fotoreceptor în bacterii - și l-am întâlnit din nou în memorie, unde schimbările în permeabilitatea membranei la nivelul sinapsei care acționează în stabilirea amintirilor umane sunt găsite funcționând în același mod în sinapsele melcilor de mare.

Modelul evolutiv ne permite, de asemenea, să înțelegem de ce există o variabilitate mai mare a trăsăturilor și funcțiilor pe măsură ce ne deplasăm de-a lungul scării filogenetice și, metaforic, în susul piramidei. Trăsăturile care evoluează mai târziu sunt mai libere, ca să spunem așa, să prezinte o variație mai mare de la individ la individ, deoarece mai puține forme sau funcții ulterioare depind de ele. Trăsăturile care nu depind neapărat de caracteristicile structurale, ci sunt legate mai mult de comportamente, sunt deschise la o serie de strategii. Căutarea hranei, reproducerea sau apărarea împotriva prădătorilor dezvăluie o multitudine de mecanisme care se dezvoltă în diferite nișe și o variație considerabilă de la individ la individ în cadrul speciei. Această variație devine o sursă de speciație pe măsură ce strategiile sau structurile noi se dovedesc a avea un avantaj adaptiv. Dar formele noi sunt întotdeauna construite pe o platformă împrumutată, ca să spunem așa, de la varietățile anterioare.

La fel se întâmplă și cu conștiința, cu mintea, cu sensibilitatea. Versiunea de bază, forma simplă pe care o găsim la baza unicelulară este consecventă între indivizi și specii. Ea prezintă funcțiile de bază pe care le-am enumerat în capitolul anterior. Pe măsură ce avansăm spre specii mai complexe din punct de vedere cognitiv și afectiv, apar o multitudine de stiluri cognitive și emoționale și de strategii funcționale. Observăm, de asemenea, o variație mai mare în cadrul speciei și între specii, iar atunci când ajungem la apoteoza conștiinței umane, descoperim o capacitate care nu se regăsește în altă parte, atât în ceea ce privește varietatea abilităților și capacităților noastre mentale, cât și variabilitatea mare de la individ la individ.

De fapt, acest decalaj între abilitățile cognitive ale hominizilor și rudele noastre cele mai apropiate este uriaș, aproape inexplicabil, mai ales în comparație cu diferențele dintre alte specii de primate. După cum am menționat mai devreme, bănuiesc că acest lucru se datorează în mare parte faptului că suntem ultimii din genul nostru, singurul membru rămas din linia hominizilor. Ceilalți care ar fi putut prezenta capacități intermediare au dispărut cu toții. Este posibil ca noi să fi jucat un rol în acest sens.

Reducționismul

În filosofie, *reducționismul* nu este o problemă simplă. Are versiuni bazate pe logică și principii abstracte, altele care restrâng discursul la unul sau altul dintre domeniile specifice de cercetare, altele care depind de faptul dacă subiectul este unul de identitate sau unul de proprietate și altele care se preocupă de modul în care sau dacă reducerea nivelurilor de descriere sau analiză produce în mod necesar un câștig epistemic. Abordarea care este relevantă pentru argumentele mele este ceea ce se numește în general reducere științifică sau materialism reductiv. Aceasta vine să ceară un răspuns la o întrebare: Se obține o înțelegere și o putere explicativă mai mare într-o anumită problemă sau domeniu de cercetare prin reducerea nivelului de analiză la un nivel mai profund, mai fundamental?

Abordarea **CBC** este extrem de *reducționistă* în acest sens. Ea se bazează pe presupunerea că ontologia originală a conștiinței poate fi, în principiu, redusă la principiile de bază ale biologiei moleculare, că rădăcinile numeroaselor varietăți de sensibilitate observate de-a lungul panoramei filogenetice se găsesc în funcțiile bio-moleculare de bază care au apărut pentru prima dată în existența speciilor unicelulare.

Evit versiunea eliminativistă a materialismului susținută de filosofii Patricia și Paul Churchland și reluată în argumentele lui Daniel Dennett. În modul de gândire al lui Churchland, conceptele care identifică entități din psihologia populară precum credința, speranța și chiar

conștiința (ceea ce în cercurile filosofice se numește atitudini propoziționale) ar trebui eliminate din discurs pe motiv că sunt lexicografice dezordonate, probabil nu corespund niciunei stări reale a naturii și nu oferă niciun câștig explicativ. Churchland și alții susțin că, odată ce detaliile neurocognitive au fost rezolvate, conștiința ca concept ontologic poate (și ar trebui) să fie eliminată din discurs. Toate problemele pot fi rezolvate prin discuții despre căile neuronale, operațiunile de transport ionic, mecanismele de permeabilitate a membranelor sinaptice, centrele cerebrale și orice mecanisme noi ar putea fi descoperite.

Există o anumită atractivitate în această poziție, în mare parte pentru că ajută la imunizarea împotriva dualismului și evită problema perspectivei la persoana întâi. Dar acest lucru are un cost și, deoarece există și alte modalități mai simple de a eluda aceste probleme - unele dintre acestea le vom analiza în secțiunea următoare - prețul filosofic cerut depășește aceste beneficii conceptuale. De asemenea, după cum am observat în discuțiile despre abordarea lui Dennett din capitolul 3, nu există nimic de câștigat prin eliminarea conceptului de conștiință. Acesta este util; se potrivește confortabil cu modul în care subiectul este discutat în general. Cel puțin, eliminarea sa din discursul convențional va conduce la unele conversații foarte greoaie.

De asemenea, este important de remarcat faptul că eliminativiștii au intrat în luptă cu accentul obișnuit pe problemele conștiinței umane, nu pe originile minții. Dacă principiul de bază pe care îl propun prin abordarea **CBC** este corect și viitorii cercetători microbiologi descoperă mecanismele specifice pentru apariția inițială a sensibilității - și mai multe despre acest lucru mai jos - natura studiilor despre conștiință se va schimba în consecință, dar limbajul utilizat nu. Teoria apariției conștiinței se va schimba, dar termenul de bază va fi în continuare util. Conștiința este reală. Statutul său ontologic este sigur. Este doar un produs al principiilor biologice.

Ar putea fi util să aruncăm o scurtă privire înapoi la o epocă anterioară, când a existat o formă diferită de gândire eliminatorie, și anume cea introdusă de *behaviorism*. În gândirea *behavioristă* timpurie, în special versiunea anunțată de John Broadus Watson în 1913[74], se susținea că constructele mentale neclare precum conștiința pot și trebuie eliminate din domeniu prin înlocuirea gândurilor cu acțiuni, a stărilor mentale cu comportamente. Poziția lui Watson a fost îmbrățișată de mulți, în mare parte pentru că deceniile anterioare au produs ceea ce era, sincer, un dezastru empiric și teoretic.

Conștiința fusese, până atunci, obiectivul principal al psihologiei experimentale, teoretice. Primii cercetători, precum Wilhelm Wundt, Carl Stumpf și Edward Titchener, erau conștienți de problema perspectivei la persoana întâi. Pentru a încerca să o eludeze, ei au folosit introspecția ca principal instrument de cercetare. Participanții la experimentele lor erau numiți „observatori” și erau antrenați să își „observe” propriile stări interne, sentimente, reacții, experiențe senzoriale, percepții și alegeri și să le raporteze în cel mai obiectiv mod posibil. Titchener, care a fost o figură dominantă în psihologia experimentală din Statele Unite timp de două decenii după sosirea sa la Universitatea Cornell în 1895, a susținut o abordare pe care a etichetat-o drept *structuralism*. Viziunea sa mai largă se baza pe speranța de a construi un fel de tabel periodic mental care să reflecte gama ordonată de elemente ale conștiinței.

În replică, James Rowland Angell de la Chicago, frustrat de ceea ce el considera a fi concentrarea îngustă a lui Titchener, a dezvoltat o abordare pe care a etichetat-o drept *funcționalism*, în care introspecția se concentra pe funcțiile și rolurile actelor mentale, nu pe structura conținutului acestora.[75] Alții au susținut că mintea nu era niciodată cu adevărat nemișcată și că aceste eforturi de a obține un fel de imagine instantanee a funcțiilor și/sau conținutului minții erau inutile; ei și-au îndreptat introspecția la persoana întâi spre raportarea proceselor mentale în curs de desfășurare, pe

74 Lucrarea era intitulată "*Psychology as the behaviorist views it*" (Psihologia așa cum o vede behavioristul). Aceasta a avut un impact puternic asupra psihologiei și a inaugurat o adevărată revoluție în domeniu.

75 Abordarea lui Angel a fost dezvoltată în cadrul psihologiei experimentale și a precedat cu câteva decenii versiunile filosofice ale funcționalismului discutate în capitolul 1. Există o suprapunere conceptuală redusă sau inexistentă cu cele câteva versiuni ale funcționalismului pe care le-am examinat în capitolul 1.

măsură ce acestea se derulau în timp. Lucrurile nu au mers bine. Descoperirile erau adesea nereplicabile, rapoartele erau confuze, limbajul utilizat pentru a descrie stările interne era dificil de înțeles, iar criticile s-au înmulțit. O mare parte a sprijinului pentru apelul lui Watson în favoarea unei științe a psihologiei obiective și riguroase a venit din ceea ce era pur și simplu un sentiment de ușurare.

Dar nici introspecția ca principiu general, separat de orice metodologie particulară, nici conștiința ca trăsătură fundamentală a minții nu au dispărut vreodată. Pentru început, conștiința era încă o componentă esențială a diverselor abordări psihoterapeutice care dominau lumea clinicienilor. Acolo, scotocirea prin mintea cuiva era nu numai utilă, ci și obligatorie. Preocupările legate de procesele subiacente ale minții nu au fost niciodată abandonate complet, nici măcar în abordările experimentale bazate mai mult pe laborator. În ciuda celor câteva decenii de hegemonie *behavioristă* în psihologia experimentală și teoretică, conștiința a continuat să fie un subiect irezistibil și și-a făcut loc în cercetarea psihofizică, în studiul percepției, al limbajului, al comportamentului de alegere și, desigur, în psihologia dezvoltării și în psihologia socială.

Odată cu diminuarea treptată a *behaviorismului* la sfârșitul anilor 1960 și în anii 1970, conștiința și mintea și-au reluat vechile locuri în centrul scenei, deși formele titcheniene și angeliene de introspecție au fost lăsate, din fericire, în mormânt. Psihologii cognitivi au continuat să trăiască cu această problemă murdară a perspectivei la persoana întâi, fără să se îngrijoreze în mod vădit de ea. După ce am lucrat în domeniu de la începutul anilor 1960 încoace, pot să vă spun că această problemă nu a apărut niciodată, cu excepția cazului în care, bineînțeles, a trecut pe acolo un filosof.

De asemenea, în mod amuzant, în acești ani au existat relativ puține eforturi pentru a încerca să definească conștiința sau să o caracterizeze într-un mod specific. Ea a devenit pur și simplu unul dintre focarele centrale ale psihologiei cognitive și ale neuroștiințelor cognitive. A fost tratată ca un obiect de studiu ontologic solid și, cu excepția cazului în care ar fi intervenit un eliminativist, puțini sau chiar niciunul dintre cei care „fac” psihologie ar fi luat în considerare renunțarea la ea în favoarea mecanismelor biologice sau neurologice subiacente. În prezent, în cadrul psihologiei, conștiința tinde să fie tratată ca un termen de „titlu de capitol” utilizat pentru a marca o secțiune dintr-un text și care, în virtutea acestui rol retoric, nu necesită definirea sau delimitarea sensului.

Așadar, să revenim la reducționismul meu. Este unul pragmatic. Atunci când vorbim despre apariția conștiinței sau despre originea minții, consider că este mai bine să ne referim la circumstanțele de bază în care aceasta a apărut pentru prima dată la cele mai simple specii, decât la cele mai înalte niveluri de dezvoltare la *Homo sapiens* sau în orice alt loc intermediar. Această mișcare nu necesită eliminarea noțiunilor abstracte și neclare precum conștiința, mintea, *qualia*, experiența fenomenală, subiectivitatea sau conștiința de sine. Toate acestea sunt concepte utile și, după cum am susținut de-a lungul timpului, trebuie încă prelucrate prin identificarea substraturilor biologice și neuronale ale fiecăruia, indiferent de specia examinată.

Încercați acest scenariu ipotetic pentru a vă face o idee despre tipul de reducționism pragmatic pe care îl favorizez. Să presupunem că un grup de oameni, majoritatea oameni de știință cu diferite specializări, stau în jurul clubului facultății universității când un coleg din programul de poezie, cu un aer destul de deprimat, intră și se prăbușește pe un scaun cu un oftat trist. Își privește prietenii morocănos și își spune povestea. A fost părăsit de soția sa, care a fugit cu un alt bărbat, un poet, nu mai puțin. Un moderator (care se întâmplă să fie președintele Departamentului de Filosofie) intră în scenă, din stânga scenei, și îi întreabă pe fiecare dintre cei adunați ce recomandă pentru a înțelege natura stării sale mentale și ce sugestii ar putea avea cu privire la acțiunile sale viitoare.

- Profesorul de finanțe îi zâmbește încurajator, menționând că, cu siguranță, nu este singurul bărbat care s-a aflat în această situație și că, în timp, va trece peste asta. Spune că toate motivele depresiei au legătură cu relațiile; în lumea lui, zâmbește cu melancolie, este vorba

întotdeauna de relații și de a afla în cine poți avea încredere. Îi sugerează să ia legătura cu familia și/sau cu prietenii săi, să iasă în oraș, să bea câteva beri, să încerce să aibă întâlniri cât mai curând posibil și, oh da, să își găsească un avocat bun specializat în divorțuri.

- Psihoterapeutul se uită nedumerit la acest lucru și arată că tipul de depresie anxioasă care însoțește acest tip de eveniment de viață provoacă schimbări în gândirea și comportamentul cuiva, iar acestea trebuie abordate. El indică datele care arată că schimbările efectuate prin TCC (terapie cognitiv-comportamentală) sunt cunoscute ca fiind utile și recomandă o serie scurtă de ședințe pentru a-i readuce gândirea, comportamentul și ideile într-o stare psihologic mai echilibrată.
- Psihanalistul dă din cap cu tristețe, spunând ceva de genul că simpla abordare a gândurilor și comportamentelor este doar superficială și că reacția sa la despărțire nu poate fi înțeleasă sau tratată decât dacă este dispus să sape până la cauzele reale ale depresiei sale. El recomandă un program complet de terapie psihanalitică care să exploreze copilăria și relațiile sale cu părinții și alte figuri de autoritate.
- Psihiatrul din sală doar râde și subliniază că toți sunt aproape jenant de simpliști. Evenimentele din
viața individului, observă el, au produs schimbări în distribuția neurotransmițătorilor din creierul său, în special dopamina și serotonina, și că, dacă nu coboară până la acest nivel și nu îi administrează antidepresive adecvate, este puțin probabil să înțeleagă ce se întâmplă în mod semnificativ. Cu siguranță, remarcă el zâmbind cu înțelepciune, nu vor înțelege altfel cauzele reale ale depresiei.
- Biofizicianul intervine apoi în discuție, arătând că psihiatrii habar nu au de ce aceste compuși funcționează - adăugând în mod parental că este posibil să nu funcționeze - iar motivul principal este lipsa unui model complet dezvoltat al acțiunilor care au loc la nivelul particulelor atomice care determină cum și de ce funcționează anumite tipuri de medicamente neurotropice și psihotrope. Aici, intonează el, este locul unde trebuie căutate cauzele epistemice ale stării sale mentale.
- Specialistul cuantic începe să deschidă gura, dar cei prezenți îl reduc la tăcere și recunosc că tipul de la finanțe ar putea avea dreptate și îl invită pe poetul încornorat să li se alăture la taverna locală pentru câteva beri.

Ideea este, desigur, că nivelul de reducere ar trebui să depindă de întrebările puse și de tipurile de răspunsuri căutate. Dacă scopul este de a-l readuce pe prietenul nostru versificator pe picioarele sale emoționale, atunci probabil că vom dori să rămânem la analiza de nivel **CBT** - deși punctele profesorului de finanțe nu sunt chiar atât de rele. Dacă dorim cu adevărat să punem stăpânire pe alte probleme, cum ar fi motivul pentru care s-a căsătorit cu această femeie, ce a determinat-o pe aceasta să se îndepărteze sau alte probleme, cum ar fi modul în care ar putea deveni mai conștient de sine și ar putea face ca viitoarele relații să funcționeze, atunci ar fi mai bine să facem o cercetare mai profundă, mai analitică. Dacă ne întrebăm ce fel de funcții cerebrale funcționează la subiectul nostru, atunci probabil că dorim să coborâm la nivel molecular și așa mai departe. Cât de mult să „reducem” este o chestiune pragmatică, una practică.

De asemenea, luăm în considerare puterea explicativă. Fiecare nivel de analiză oferă o perspectivă asupra caracteristicilor cauzale care operează la acel nivel. Dacă sunteți un terapeut specializat în **CBT**, atunci aici opriți reducerea, pentru că aici obțineți tipurile de răspunsuri la tipurile de întrebări pe care le puneți și aici este cel mai probabil să găsiți o explicație satisfăcătoare a stării mentale actuale și viitoare a individului. Dar dacă vă specializați în structura moleculară subiacentă a elementelor care alcătuiesc neurohormonii din creierul uman, acest nivel cognitiv-comportamental nu vă este de mare ajutor. Veți pune întrebări diferite și veți aștepta tipuri diferite de răspunsuri și veți căuta niveluri diferite de reducere pentru acestea. O putere explicativă mai mare nu este neapărat obținută prin exerciții reducționiste. Depinde de ceea ce încercați să explicați.

În cadrul poziției stabilite de cadrul **CBC**, întrebările ar trebui să fie cele puse cu privire la efectele materialiste care cauzează apariția sensibilității, iar nivelul la care trebuie să ne reducem

este cel în care factorii bio-moleculari, genetici, operează în speciile unicelulare. Desigur, este foarte posibil - de fapt, mai mult decât posibil, este aproape sigur - că se vor găsi niveluri cauzale mai profunde bazate pe principii biochimice, fizice. Este probabil ca structura moleculară a materialelor care se combină pentru a produce primele stări subiective să poată fi ele însele mai bine înțelese prin examinarea componentelor lor elementare. Dacă va fi util să coborâm la acest nivel și să schimbăm modul în care sunt discutate probleme precum conștiința și mintea va fi o chestiune de practică, comunicare, explicație cauzală și alți factori precum parcimonie, eleganță și natura întrebărilor puse.

Din nou, mă aflu în mare parte de acord cu John Searle, care a comentat de mai multe ori această problemă. Conștiința, în opinia sa, este tocmai rezultatul anumitor tipuri de acțiuni bio-moleculare. Pentru a folosi câteva dintre exemplele sale preferate, pe care le-a prezentat pentru a sublinia acest lucru: fotosinteza este reală în acest fel, la fel ca respirația, digestia și chiar viața însăși. Toate acestea sunt produse ale unor funcții materiale, iar cele mai eficiente și practice teorii ale lor s-au bazat pe principii bio-chimice, fiziologice și psihologice. Am făcut o treabă destul de bună în ceea ce privește detaliile fotosintezei, respirației și digestiei (nu la fel de bine în ceea ce privește viața), dar încă câștigăm atât în ușurința comunicării, cât și în înțelegerea funcției, dacă ne agățăm de conceptele generice, în loc să le eliminăm. Chiar și atunci când vom descompune evenimentele criteriale de la nivel molecular care generează sensibilitatea, ne vom confrunta în continuare cu sarcina descurajantă de a înțelege mecanismele evolutive care acționează atunci când aceste stări și procese mentale sunt modificate pe măsură ce specii noi se ramifică.

Reducerea originii conștiinței la rădăcinile sale în organismele simple, unicelulare, ne va aduce și unele avantaje secundare. În primul rând, aceasta va schimba multe dintre întrebările puse - și, după cum s-a observat de multe ori, o astfel de schimbare în cercetarea epistemică poate fi deschiderea către găsirea unor răspunsuri mai bune. Celelalte aspecte care au fost în centrul cercetării nu sunt afectate, însă cadrul general în care se desfășoară acestea se va schimba de la „Cum face creierul mintea?” la „Cum formează sistemele organice de complexitate variabilă baza pentru forme mai sofisticate de conștiință și stări mentale”.

Evitarea misterianismului

Eticheta „Noii misteriști” a fost inventată de filosoful Owen Flanagan de la Universitatea Duke în influenta sa carte „*Știința minții*” (The Science of the Mind). Ea este folosită în sens larg pentru acei filosofi care susțin că *problema dificilă* a conștiinței este iremediabil refractară la eforturile de găsire a unei soluții și, ca urmare, s-au retras în poziții pe care filosofii cu o înclinație materialistă le consideră neștiințifice. Acestea se prezintă sub două forme generale: echipa „este greu de spart”, care urmează exemplul lui Thomas Nagel, susținând că cunoașterea cauzelor materiale nu ne va oferi niciodată vreo apreciere a experienței individuale, subiective, și echipa „este o nucă de spart”, formată din gânditori precum Colin McGinn, care susține că pur și simplu nu suntem echipați intelectual pentru a reuși acest lucru. În opinia lui McGinn, suntem „excluși” de la a înțelege vreodată orice soluție posibilă la *problema dificilă*. Marțienii ar putea să o înțeleagă, dar noi nu am înțelege niciodată teoria dacă ne-ar pune-o la dispoziție.

Filozofii „intransigenți” susțin că este probabil imposibil să aflăm cum creierul creează mintea. Indiferent cât de mult am învăța despre relațiile dintre funcțiile și structurile neuronale ale creierului și procesele comportamentale și cognitive la care acestea sunt legate, tot nu vom reuși să înțelegem *qualia* sau stările mentale. Această poziție, după cum am menționat în mai multe locuri, a apărut din argumentul lui Nagel „cum este să fii” și a fost extinsă pentru a încorpora poziția conform căreia chiar și o descriere materialistă completă a materialului nu va reuși să explice

mentalul. Un exemplu favorit este cel al savantului daltonist care știe tot ce se poate ști despre biofizica și neurofiziologia vederii culorilor pe unde scurte, dar care încă nu are niciun sentiment fenomenal despre cum arată albastrul.

Problema mea cu acest argument, acum celebru, este că, din punct de vedere științific, pur și simplu nu este relevant. De asemenea, nu știm cum este să fii o specie fotosintetizantă, dar acest lucru nu ne-a împiedicat să dobândim o înțelegere profundă a bazei biochimice a procesului de colectare a energiei din lumina soarelui. Referirea constantă la „golul explicativ" care rezultă din necesitatea de a adopta „perspectiva primei persoane" poate fi, și judecând după numărul de lucrări scrise pe această temă, cu siguranță este, un subiect demn de dezbatere în rândul filosofilor, dar nu ar trebui să fie considerat un impediment în calea progreselor în domeniul neuroștiințelor cognitive. Nu ar trebui să fim îngrijorați de faptul că nu suntem capabili să apreciem subiectiv cum arată lumea prin fascinanții ochi compuși ai insectelor și crabilor potcoavă. Ar trebui să ne bucurăm că am ajuns să știm multe despre modul în care funcționează, cum și când au evoluat și ce caracteristici au în comun cu alte specii care au dezvoltat alte moduri de procesare a luminii. Da, sună arogant, dar răspunsul pe termen lung la clasica enigmă Nagel ar putea fi „și ce dacă".

Al doilea grup sunt adevărații „disperați", cei care susțin că *problema dificilă* nu poate fi rezolvată niciodată, cel puțin nu de către noi. McGinn, care a fost cel mai influent susținător al acestei poziții, nu se opune existenței unui set de principii cauzale, astfel încât substanțele materiale neurofiziologice să producă experiența subiectivă. El susține pur și simplu că nu suntem echipați din punct de vedere cognitiv pentru a ști care sunt acestea sau pentru a le înțelege. Există lucruri pe care un cimpanzeu le știe sau le poate înțelege și care nu sunt la îndemâna unui ornitorinc. Există lucruri pe care le știm și le înțelegem și care sunt iremediabil în afara posibilităților cognitive ale unui cimpanzeu. Și, urmând lanțul logic al raționamentului, există lucruri pe care pur și simplu nu le putem cunoaște sau înțelege, cum ar fi reprezentarea spațiului cu 10 dimensiuni. McGinn consideră că există aproape sigur un răspuns la *problema dificilă*, dar acesta este la fel de inaccesibil pentru noi cum este algebra pentru un bonobo.

Atunci când Flanagan a introdus eticheta, el a pus în contrast acești misteriști „noi" (sau epistemologici) cu misteriștii anteriori, „vechi" (sau ontologici). Versiunea mai nouă, cea care ne interesează aici, este marcată de un angajament ferm față de naturalism, credința că conștiința și mintea sunt reale și au la bază mecanisme biofizice de bază. Nu există elemente non-naturale aici, nici suflet, nici spirit, nimic evanescent sau nonmaterial. Problema nu este în ființă, ci în cunoașterea despre ființă. Pe scurt, niciuna dintre cele mai noi abordări nu adoptă o ontologie dualistă în care mintea și corpul sunt privite ca domenii separate care urmează principii și legi distincte. Poziția lor este că *problema dificilă* este foarte, foarte dificilă, atât de dificilă încât fie este puțin probabil să o rezolvăm vreodată într-un mod care să fie satisfăcător din punct de vedere epistemic, fie pur și simplu nu vom fi niciodată capabili să înțelegem răspunsul - chiar dacă ne-ar fi prezentat.

Flanagan a opinat, de asemenea, că acestea nu sunt argumente nerezonabile - deși nu este de acord cu ele. Și eu sunt de acord. În primul rând, pur și simplu nu sunt de acord cu afirmația că, doar pentru că nu am rezolvat o *problemă dificilă*, fie nu o vom rezolva, fie suntem o specie prea proastă pentru a o rezolva. Singurul lucru care poate fi câștigat aruncând prosopul aici este câștigarea unui argument filosofic, iar asta, în felul meu de a gândi, este o victorie destul de neimpresionantă.

Luați oricare dintre numeroasele domenii de cercetare extrem de abstracte, cum ar fi mecanica cuantică sau topologia, astrofizica, neurobiologia, orice altceva. Urcați în mașina timpului, întoarceți-vă în timp și prezentați cadrul conceptual al mecanicii cuantice, al astrofizicii sau al oricăruia dintre celelalte în fața ochilor neîncrezători ai celor mai incisivi gânditori din secolul al XVIII-lea. Ați obține aceleași tipuri de nedumerire reflectate în ochii noilor misteriști de astăzi. Dar am înțeles mecanica cuantică - sau cel puțin o mare parte din ea. Obținem în mod regulat informații noi despre principiile astrofizice care caracterizează universul nostru. În plus, teoreticienii cuantici și astrofizicienii se simt destul de confortabil conceptualizând principiile subiacente - chiar dacă înțelegerea lor epistemică este dependentă de matematică. Nu văd nimic de

câştigat şi multe de pierdut prin afirmaţiile că nu vom fi niciodată capabili să ne dăm seama cum creierele creează minţile. Şi, desigur, bănuiesc că a începe cu modul în care procariotele au reuşit acest truc va fi un punct de plecare destul de solid.

Mai rău, este clar că o implicaţie a misterianismului este că ne cere să ne retragem din laboratoare, să ne retragem din teorie şi din construirea de modele. Această invitaţie ciudată la insularitate a declanşat amintirea unei parabole atribuite lui Francis Bacon despre grupul de teologi învăţaţi care se luptau să rezolve problema numărului de dinţi din gura unui cal. După multe zile de reflecţii şi dezbateri asupra semnificaţiilor şi interpretărilor Scripturii şi ale textului, ei nu găsesc nicio soluţie şi stau disperaţi. Un frate tânăr intervine şi le sugerează să încerce să se uite în gura calului şi să numere dinţii. Acest sfat profan este întâmpinat cu groază de fraţi care, în cuvintele lui Bacon, s-au înfuriat foarte tare; şi, unindu-se într-un tumult puternic, au năvălit asupra lui şi l-au lovit, la şold şi la coapsă, şi l-au alungat imediat .... [şi] ei, ca un singur om, declarând că problema este o taină veşnică... au ordonat ca aceasta să fie consemnată.

Oricât de amuzantă ar fi această mică ironie la adresa dogmaticilor din biserică, există mai mult în spatele ei decât pare. Din punct de vedere teologic al preoţilor, poziţia lor antiempirică are de fapt sens. În sistemul lor de credinţă, Dumnezeu a făcut lumea şi a creat toate creaturile care o populează. El a avut desenele şi planurile sale pentru fiecare dintre creaturile sale şi le-a creat în consecinţă. Ceea ce ştiu ei despre marele său proiect se găseşte, de fapt, se poate găsi doar în textele sacre, deoarece acestea sunt vehiculele pe care Dumnezeu le-a folosit pentru a-i informa despre ceea ce a pus la cale. A şti câţi dinţi are un cal înseamnă, prin urmare, a şti ceva despre planul lui Dumnezeu şi, pentru că scripturile şi textele şi interpretările lor de către cercetători aprobaţi oferă cea mai bună perspectivă asupra acestui plan, acestea sunt locurile unde se va găsi adevărul.

Şi devine şi mai bine (sau mai rău). Deoarece Satana se joacă mereu cu lucrurile - cum ar fi încurajarea tinerilor călugări să se frece de eretici - nu există niciun motiv pentru ei să creadă că numărarea dinţilor ar rezolva puzzle-ul. Diavolul ar fi putut cu uşurinţă să scoată câţiva dinţi sau să facă să crească dinţi supranumerari. Calul ar fi putut suferi un accident şi să piardă unul sau mai mulţi dinţi. Preotul care făcea numărătoarea ar fi putut face o greşeală. Orice număr de măsuri empirice ar fi putut să îi inducă în eroare. Aşa că, la fel ca în cazul lui McGinn, au fost de acord că, în cadrul viziunii lor asupra lumii, puzzle-ul era de nerezolvat, de neştiut pentru mintea umană.

În cadrul **CBC**, argumentele noilor misterioşi epistemologici nu sunt deosebit de relevante. Putem depăşi obiecţiile lui Nagel şi McGinn pur şi simplu ignorându-le şi făcând ştiinţă bună la nivel microbiologic, deoarece aceasta ne invită să adoptăm o viziune diferită asupra lumii. Cu toate acestea, ceilalţi misterişti, cei ale căror poziţii provin din perspective ontologice mai fundamentale, cum ar fi *dualiştii* şi *panpsihicii*, creează probleme - în special deoarece au fost îmbrăţişaţi de mai mulţi cercetători contemporani. Ne-am ocupat de panpsihism în capitolul 2. Merită să aruncăm o privire asupra dualismului şi a motivelor pentru care unii ca David Chalmers l-au susţinut şi de ce, din perspectiva **CBC**, nu este o poziţie utilă.

Un scurt comentariu despre dualism

Ca aproape orice poziţie din filosofie, dualismul are mai multe versiuni. Toate au în centrul lor convingerea metafizică că realitatea este alcătuită din două tipuri distincte de „lucruri”. Modelele clasice, mai vechi, împart realitatea în material şi imaterial; cele cu o tentă teologică disting corpul de suflet; iar versiunile actuale, seculare, pun punctul de clivaj între material şi mental. Unii dintre aceşti ultimi dualişti, cum ar fi paraleliştii, încearcă să nu se rătăcească în grotele misteriene. Calea lor de scăpare este de a presupune că materialul şi mentalul sunt două versiuni ale realităţii cu o legătură, în principiu, de cunoscut între ele. Orice evenimente care au loc în lumea fizică produc stimuli care, atunci când impactează sistemele senzoriale umane, provoacă

experienţe subiective. Aceştia sunt cunoscuţi şi sub denumirea de *dualişti psihofizici*, deoarece ei consideră că există o legătură între evenimentele fizice şi experienţa psihologică a acestor evenimente. Pentru paralelişti, punctul sensibil este colectarea datelor corecte şi evitarea tipurilor de probleme care i-au afectat pe introspecţionişti.

Problema pentru filosofia minţii şi, prin extensie, pentru neuroştiinţele cognitive, vine din partea celor care îmbrăţişează propunerea conform căreia experienţele fenomenale pe care le avem se află pe un plan ontologic distinct de realitatea obiectivă, fizică, care poate fi înţeleasă în termeni pur materialişti. Această poziţie este, din punctul meu de vedere, absurdă şi antitetică oricărei poziţii ştiinţifice coerente. De asemenea, este contraproductivă în sensul că adoptarea ei înseamnă, ca şi poziţia lui McGinn, fluturarea steagului alb al capitulării şi plecarea în faţa ignoranţei.

Din nou, ne putem întoarce la Searle, care a furnizat unele dintre cele mai bune argumente împotriva oricărei versiuni a dualismului, în special prin sublinierea faptului că distincţiile reale nu sunt cele dintre minte şi materie, ci între acele caracteristici ale lumii care există în virtutea proprietăţilor lor fizice şi cele care necesită un observator. Astfel, electricitatea, scaunele, gravitaţia, copacii sunt din primul tip, în timp ce banii, legile, societatea şi minţile sunt din al doilea.

Dar Searle, în respingerea dualismului minte-corp, respinge şi materialismul. Şi aici este unde ne despărţim. În primul rând, apreciez că nu am nicio problemă cu poziţia lui Searle conform căreia materialismul, ca şi reducţionismul, te poate trage în găuri de iepure eliminativiste. Acest lucru se întâmplă, după cum am menţionat, din cauza prescripţiei de a reformula discursul, astfel încât discuţia să fie despre neuroni şi centre corticale, în loc să fie despre stări conştiente şi experienţe mentale. Searle preferă, de asemenea, şi din nou sunt de acord, să trateze conştiinţa ca pe un rezultat subiectiv, ontologic sigur, al anumitor procese biologice, ca pe un lucru care este privit dintr-o perspectivă de la prima persoană. Dar problema mea cu Searle, aşa cum am observat la începutul acestei călătorii, este că el, la fel ca aproape toţi filosofii care au îndrăznit vreodată să încerce să descifreze misterele conştiinţei, îşi începe (şi îşi încheie) lupta cu creierul uman şi cu viaţa mentală umană.

Din perspectiva **CBC**, materialismul şi reducţionismul nu sunt doar poziţii legitime, ci şi obligatorii. Dacă originile minţii se găsesc în procesele biochimice şi moleculare fundamentale ale celei mai simple specii unicelulare, atunci este necesar să coborâm până la acest nivel şi să ne dăm seama care sunt aceste mecanisme şi cum se instanţează ele într-un organism atât de primitiv. Următoarea serie de paşi ar fi să urmărim evoluţia acestei conştiinţe până la versiunea care a evoluat în noi. La fiecare etapă a procesului, mecanismele ar fi mai complexe, dar toate s-ar baza în continuare pe principii fizice, ştiinţifice. Dacă Searle şi alţii care adoptă poziţia sa filozofică doresc să continue discuţiile despre aceste chestiuni axate pe *Homo sapiens* folosind perspectiva la persoana întâi, este în regulă. De fapt, atunci când subiectul discursului este mintea umană, aproape sigur va continua să fie cel mai simplu şi practic mod de a face ştiinţă a conştiinţei.

Revenind la apărarea mea a lexicografiei dezordonate, există motive pentru care nu am încercat să definesc în mod obiectiv termeni precum conştiinţă, minte, fenomenal sau sensibilitate. Există motive pentru care am prezentat adesea argumente folosind semnificaţiile psihologice populare ale termenilor şi principiilor. Există motive pentru care nu am încercat să disec fiecare dintre perspectivele rafinate şi bogate din punct de vedere filosofic care se regăsesc în literatura de specialitate. Aceasta se datorează faptului că abordarea **CBC** insistă asupra unei poziţii mecaniciste cu privire la originile sensibilităţii, apreciind în acelaşi timp că, atunci când vine vorba de scena centrată pe om pe care păşesc practic toţi filosofii, intervin chestiuni practice. Mintea unei bacterii şi cea a unui om nu sunt decât poluri pe continuumul conştiinţei. Trebuie să putem vorbi despre ele în moduri care să ne permită să facem ştiinţă bună şi să descoperim mecanismele biologice, evolutive şi psihologice care acţionează, fără a ne împotmoli în dispute lexicografice.

Probleme empirice în căutare de soluții

Există câteva(!?) alte preocupări care însoțesc **CBC**. Să facem o scurtă trecere în revistă a acestora. Aceste discuții se vor axa pe probleme și întrebări, nu pe soluții sau răspunsuri - deși vom încerca să aflăm pe ici pe colo unde sau cum ar putea fi găsite aceste soluții și răspunsuri.

Care sunt procesele moleculare care stau la baza sensibilității?

Sincer, habar nu am. Nimeni nu știe pentru că, din câte am putut constata, nimeni nu a încercat să le descifreze - în sensul de a face un efort empiric în laborator. Cu toate acestea, câteva încercări, în mare parte de natură speculativă, identifică mecanisme care sunt cel puțin plauzibile - iar eu voi prezenta ipoteza mea candidat, care va fi o speculație pură, dar potențial utilă. De asemenea, voi prezenta modalitățile prin care cercetările viitoare privind genetica speciilor unicelulare ar putea oferi un sprijin empiric (sau, desigur, o falsificare) a acestor procese.

Propunerea Sacks et al.

Am întâlnit deja acest efort în capitolul 2, când am examinat alte abordări ale problemelor din biologia celulară care reflectau unele dintre caracteristicile modelului **CBC**. Este vorba despre scrisoarea deschisă adresată lui Christof Koch de către un grup de neurocercetători importanți[76] (Oliver Sacks, Antonio Damasio, Gil Carvalho, Norman Cook și Harry Hunt). Aceștia au subliniat obiectivul lor care, după cum s-a menționat, este similar cu cel pe care îl dezvoltăm aici.

Căutarea proprietăților care stau la baza conștiinței până la nivelul protozoarului (un organism unicelular) pentru a identifica mecanismele fundamentale la nivel celular care, atunci când sunt extinse în sisteme nervoase complexe, dau naștere proprietăților care sunt denumite în mod obișnuit „minte”.

Diavolul, desigur, stă în detalii. Sacks și colegii săi încă se agață de ideea că „mințile” (deși ei nu specifică la ce specie se gândesc aici) necesită un sistem nervos, însă rădăcinile conștiinței se vor găsi la nivel celular. Candidatul lor pentru mecanismul care guvernează această apariție a subiectivității este „excitabilitatea” sau „iritabilitatea” membranei celulare, modulată prin deschiderea și închiderea porilor care permit schimbul de ioni. Operațiunea funcțională, în opinia lor, are loc atunci când cationii (ioni încărcați pozitiv) trec în interiorul alcalin al celulei. Interesant este faptul că ei se referă la acest proces ca la „ ... *fenomenul cheie implicat în «conștientizarea» de către celulă a mediului său (”sensibilitate„)*”. Și da, termenii din ghilimelele de speriat provin direct din scrisoarea lor. Cook, Carvalho și Damasio au aprofundat această poziție într-o lucrare din 2014, în care au adus mai multe dovezi și argumente cu privire la rolul excitabilității membranei în stabilirea subiectivității, a sensibilității la speciile unicelulare.

Este important de remarcat faptul că acest proces, prin care celulele reacționează prin schimbul de ioni încărcați printr-o membrană, care a fost unul dintre pilonii modelului standard al biologiei celulare încă de la primele lucrări ale lui Alan Hodgkin și Andrew Huxley în anii 1930, este același mecanism în cele trei tipuri de celule primare găsite la animale - celulele senzoriale, celulele musculare și neuronii. Există trei motive pentru a ne concentra asupra acestui mecanism ca fiind un candidat legitim pentru mecanismul operațional comun:

- În primul rând, arată clar că sensibilitatea este un fenomen biologic și nu va fi observată în sistemele anorganice. Această concluzie este, desigur, aceeași la care am ajuns în capitolul 1.
- În al doilea rând, identifică un mecanism potențial care poate fi examinat experimental și deschide calea pentru viitorii cercetători, deoarece, după cum spun Sacks și colegii săi,

76 Damasio m-a informat (comunicare privată) că există întrebări cu privire la proveniența exactă a scrisorii și cine a scris-o. Cu toate acestea, punctele din scrisoare sunt relevante independent de această problemă.

Pentru a produce „conştiinţa” de nivel superior a organismelor animale, activitatea acestor numeroase celule excitabile pentru a obține un fel de sensibilitate trebuie să fie sincronizată (în moduri care urmează încă să fie determinate) pentru un comportament coerent la nivel de organism.

Acest comentariu atinge, dar nu abordează pe deplin, o problemă importantă cu care trebuie să se confrunte orice model care începe la nivel celular: cum se ajunge de la sensibilitatea lumii unicelulare la cea a lumii multicelulare? Ce operațiuni permit ca răspunsurile subiective separate ale fiecăreia dintre celulele individuale din „*gestalt*” să fie subsumate întregului? După cum remarcă Sacks și colegii săi, aceasta este o situație „**TBD**” clasică. Voi avea ceva mai multe de spus despre aceasta mai târziu în acest capitol, deoarece este un domeniu de cercetare intens urmărit în biologia celulară.

- În al treilea rând, observați cuvântul „animal” din citatul de mai sus. Sacks și colegii săi pun bazele unei distincții între domeniul faunei, distinct de cel al florei (și al ciupercilor). Implicația acestei notări trecătoare a lui Sacks și a colegilor săi este că ei nu cred că plantele au forma primară de sensibilitate pe care o au animalele. Tind să fiu de acord cu această poziție, dar, încă o dată, a avea doar o opinie nu înseamnă mai mult decât a avea doar o opinie. Problema va fi analizată mai îndeaproape mai jos. Vom vedea că există mai mulți botaniști și biologi celulari care susțin că plantele sunt sensibile.

Ceea ce Sacks și compania au prezentat este un scenariu rezonabil pentru modul în care a apărut prima noastră *simțire*. Este un candidat respectabil, deoarece oferă unui organism unicelular procesele necesare pentru a naviga în lumea înșelătoare a celor foarte mici. Este un mecanism care a demonstrat o mare generalitate, deoarece a stat la baza funcțiilor celulare într-o multitudine de alte tipuri de celule dintr-o mare varietate de specii. Și, ceea ce este important, pare a fi un mecanism care ar putea funcționa într-o manieră care să permită celulelor individuale să renunțe la propria lor sensibilitate endogenă pentru conştiinţa colectivă unitară care este observată în lumea pluricelulară.

<u>Sugestia lui Dennis Bray</u>.

Singura altă propunere pe care am întâlnit-o cu privire la modul în care ar putea apărea *sensibilitatea* la organismele unicelulare a fost cea prezentată de microbiologul Dennis Bray. Sugestia lui Bray nu a fost o propunere directă a unui mod în care speciile unicelulare ar putea dezvolta sensibilitate. Aceasta făcea parte dintr-un eseu scris de el în 2011, în care compara procesele organice pe care un agent biologic simplu le utilizează pentru a răspunde la schimbările de temperatură ambientală cu cele ale unui artefact, un termostat fabricat. În acest fel, a prezentat și mecanismele biologice prin care o specie unicelulară ar fi putut dezvolta un simț subiectiv al temperaturii. O extrapolare simplă a acestui proces la alte sisteme senzoriale, motorii și emoționale, împreună cu analiza lui Bray, oferă un cadru plauzibil în care ar putea apărea o sensibilitate ontologică de tipul celei descrise în capitolul anterior.

Bray și-a început explorarea observând asemănările superficiale dintre un artefact, un termostat produs de om, și un proces biologic, funcțiile termotaxice ale unei bacterii. Ambele au un mecanism comun în sensul că răspund la o gamă de temperaturi și, atunci când sunt atinse anumite puncte critice, inițiază procese care restabilesc nivelurile optime de temperatură. Alegerea controlului termic ca dispozitiv retoric este una interesantă - nu puțini neuroștiințifici și filozofi care au în vedere posibilitatea unor artefacte sensibile (pe care le-am explorat în capitolul 1) folosesc termostatele ca exemple de entități artificiale care ar putea manifesta o formă de sensibilitate nebiologică.

Cu toate acestea, după cum observă Bray, deși operațiunile sistemului anorganic pot avea paralele cu cele organice, există o diferență evidentă: artefactul nu evoluează. Este ceea ce este și nu va dezvolta operațiuni mai sofisticate fără intervenție. Sistemul termotaxic din, să zicem, *E. coli* a evoluat de-a lungul generațiilor și, treptat, s-a stabilit un simț al „cunoașterii” (și, da, Bray scrie despre cunoaștere și cunoștințe atunci când se referă la starea internă a unei bacterii) cu privire la temperatură, gradienții de temperatură și cum și când să se miște pentru a optimiza funcțiile metabolice celulare interne.

Acest lucru se întâmplă într-un mod destul de apropiat de descrierile făcute în capitolul 4 cu privire la modul în care organismele unicelulare răspund la schimbările în compoziția moleculară a mediului lor. Așa cum descrie Bray procesul, acesta începe cu:

[Un] grup de proteine, chemoreceptori, în membrana bacteriană [care]... pot exista în diferite forme sau conformații, fiecare cu un aranjament distinctiv al atomilor [care]... se modifică cu temperatura. În funcție de ... receptor și de cât de mult a fost modificat în celulă... acesta poate fi activat de căldură, activat de frig sau insensibil la temperatură.

Mecanismul complet, adaptiv, funcționează astfel cum am văzut în capitolul anterior. Dacă chemoreceptorii detectează că organismul s-a deplasat într-o regiune prea caldă (sau prea rece), acesta își va roti flagelii în sensul acelor de ceasornic și se va angaja în schimbări aleatorii de direcție (numite „rostogolire”) pentru a se îndepărta de condițiile neplăcute. Dacă temperatura detectată la capătul său este acceptabilă, flagelul își schimbă rotația flagelilor în sensul invers acelor de ceasornic și înaintează. „Comutatorul” este punctul în care receptorii de temperatură interacționează cu proteinele interne care semnalează motoarelor flagelare să se rotească într-un sens sau altul. Este puțin mai complicat decât atât, desigur, deoarece există alte straturi ale procesului care implică cooperarea kinazelor pentru a cataliza operațiunile și un număr de proteine specifice care stimulează aceste acțiuni, dar principiul de bază este cel descris.

În plus, atunci când fac parte dintr-un colectiv, modul în care bacteriile utilizează acest mecanism termostatic este sensibil la densitatea coloniei. De exemplu, în timp ce preferința implicită a bacteriilor individuale este un mediu cald, de până la aproximativ 30 °C, acestea se vor deplasa spre un loc mai răcoros de până la 18 °C dacă densitatea populației crește. Mecanismul de aici utilizează funcțiile *quorum sensing* pe care le-am discutat anterior și este foarte adaptiv. Densitatea ridicată a populației produce de obicei o reducere a nivelului mediu de nutrienți pentru fiecare individ. Atunci când sursele de hrană sunt reduse, devine avantajoasă reducerea funcțiilor metabolice și încetinirea reproducerii - procese care se desfășoară cel mai bine într-un mediu mai răcoros.

Acest tip de operațiune homeostatică, așa cum am văzut mai devreme, este bine cunoscut. Pasul pe care Bray îl face în continuare este cel interesant. El aduce în discuție procesele biologice evolutive pentru a sublinia faptul că, pe măsură ce fiecare dintre aceste funcții s-a stabilit, ea a devenit codificată în ADN-ul speciei. În timp, acumularea de moduri de a reacționa la lumea exterioară, abilitatea delicată de a detecta compoziția moleculară a mediului și de a o compara cu condițiile endogene interne și cu stările metabolice, multitudinea de sisteme senzoriale, gama nebună de proteine care au fost cooptate în aceste funcții au condus inexorabil la apariția sensibilității. Cu privire la modul în care acest lucru s-a întâmplat, în cuvintele sale,

Primul pas ar fi putut fi mutațiile care au instalat din întâmplare o secvență de aminoacizi care a produs o schimbare deosebit de mare în structura unei anumite proteine în funcție de temperatură. Ar fi urmat apoi modificări ale altor proteine care le-ar fi permis să interacționeze și să răspundă la proteina sensibilă la temperatură. În cele din urmă, ar fi putut fi construit un lanț complicat de cauzalitate biochimică prin care procesele influențate de temperatura ambientală, cum ar fi motilitatea sau metabolismul.

Nu numai că este o imagine plauzibilă, dar a fost una care ar fi caracterizat modul în care aceste organisme primitive se confruntau cu alte aspecte ale mediului lor, lumina soarelui, presiunea, compoziția moleculară a apei de mare. Același tip de mecanisme evolutive ar fi funcționat pentru toate celelalte funcții senzoriale, perceptive, motorii, afective și comunicative ale acestor organisme unicelulare. De fapt, compilația completă a acestora, luată ca o reprezentare mentală holistică coordonată, reprezintă ceea ce am numit conștiință la aceste specii primitive.

Bray nu o spune niciodată chiar așa, dar dacă citesc ceea ce susține, s-ar părea că fiecare dintre aceste mecanisme senzoriale a oferit un avantaj suplimentar în lupta pentru supraviețuire și fiecare a contribuit la un sentiment tot mai mare al unui sine fenomenal, un sine experiențial. Fiecare dintre acestea a adăugat o altă fărâmă de sensibilitate și fiecare a sporit șansele de supraviețuire. Dacă fiecare dintre aceste funcții a evoluat separat și a fost codificată treptat, una câte una, în ADN-ul speciei sau dacă toate au ajuns, ca să spunem așa, la formarea inițială a vieții este încă o întrebare deschisă - deși una la care ar putea răspunde viitoarele analize genetice.

În cele din urmă, în acest eseu, Bray face un comentariu care reflectă o mare parte din mesajul prezentat de modelul **CBC**:

Dar ce cuvinte putem folosi pentru a descrie aceste informații celulare? Biologia contemporană îmbrățișează reducționismul și evită vitalismul. Ea a avut un succes extraordinar în dezvăluirea structurilor și funcțiilor moleculelor biologice, adesea la scară atomică. Ne-a lăsat însă cu o respingere extremă, aproape puritană, a oricărei descrieri a proceselor biologice care merge dincolo de fizică sau chimie. Acesta este motivul pentru care un termen precum „cunoaștere" stârnește antipatie atunci când este aplicat celulelor sau organismelor simple - deoarece este încărcat cu conexiuni umane. (Subliniere adăugată.)

O întrebare care merită pusă aici este dacă aceste două scenarii speculative, ipoteza iritabilității lui Sacks și colegilor și evoluția fragmentată a funcțiilor specifice a lui Bray, sunt compatibile între ele. Deoarece fiecare este speculativ și niciunul nu a elaborat detalii biomoleculare specifice sau nu a efectuat nicio activitate empirică, este greu de știut. Dar, la prima vedere, ele nu sunt (sau nu par a fi) contradictorii între ele. Tipurile de mecanisme de permeabilitate a membranei celulare identificate de grupul lui Sacks fac parte din abordarea principală a microbiologiei, iar abordarea lui Bray este fundamentată pe principiile de bază ale biologiei evolutive. Așadar, avem de fapt două modele care se concentrează pe mecanisme distincte, dar utilizează principii de bază similare.

Ofertele lui Arthur Reber.

Doresc să ofer câteva sugestii cu privire la tipurile de cercetare bio-moleculară care ar putea descoperi factorii genetici responsabili pentru prima apariție a minții. Deși speculativă, următoarea schiță se bazează pe două progrese științifice recente care, în principiu, pot fi utilizate ca instrumente empirice pentru a testa implicațiile modelului **CBC**:

- tehnici de editare a genelor precum CRISPR
- algoritmi de învățare profundă în bioinformatică

Să începem cu viziunea lui Bray asupra modului în care funcțiile și operațiunile specifice au apărut și au evoluat pentru a vedea cum abordarea CBC ar putea conduce la o înțelegere a fundamentelor sensibilității la procariote. El face referire la ADN-ul bacterian. Narațiunea sa este o poveste despre funcțiile senzoriale, dar poate fi exprimată ușor diferit, astfel încât să devină o poveste despre funcționarea mentală a bacteriei. Sugestia mea este că acesta ar putea fi exact modul în care anumite segmente din ADN-ul procariotelor produc conștiința, utilizând aceleași mecanisme pe care le-a identificat Bray. În loc să moduleze și să controleze diferitele funcții perceptive și

afective care alcătuiesc conținutul minții bacteriene, aceste operațiuni ar funcționa, de asemenea, pentru a instanția experiența subiectivă internă care, după cum am susținut, este o caracteristică esențială a oricărui organism mobil.

Bacteriile se pricep să își modifice ADN-ul. Acesta este, așa cum am văzut în capitolul anterior, un mecanism care permite formarea memoriei genetice. De asemenea, acesta este modul în care își dezvoltă imunitatea împotriva virusurilor sau plasmidelor invadatoare. Ceea ce susțin eu este că, atunci când viața a început și când construcția formală a ADN-ului a reușit să scrie, ca să spunem așa, scenariul unui organism viu, autoreplicant, acesta conținea secvențe care codificau sensibilitatea. Aceste secvențe ar fi, dacă această imagine are valoare, încă prezente în ADN-ul fiecărei specii vii (cu posibila excepție a plantelor - mai multe despre acest aspect mai târziu în acest capitol).

Una dintre caracteristicile intrigante ale acestei ipoteze este că, în principiu, ar putea fi explorată experimental folosind tehnicile care permit geneticienilor să editeze ADN-ul unui organism prin tăierea unor fragmente specifice de ADN, mutarea lor în altă parte a genomului și inserarea altor fragmente de material genetic.[77] Odată identificate gena sau genele candidate pentru conștiință, rolurile lor specifice pot fi investigate prin ștergerea lor, inserarea altor secvențe în locul lor, mutarea lor în altă locație și observarea consecințelor.

Următoarea întrebare, desigur, este cum să identificăm genele țintă. Progresele recente în domeniul informaticii care utilizează tehnici de învățare profundă sunt, în principiu, capabile să îndeplinească aceste sarcini. Algoritmii de calcul moderni sunt foarte buni la descoperirea covariațiilor dintre elementele îngropate în medii extrem de complexe, cum ar fi cele din biochimia unor lungi șiruri de ADN. Aceștia sunt utilizați în prezent în bioinformatică pentru a identifica genele responsabile de diverse afecțiuni. Grupul Olga Troyanskaya de la Universitatea Princeton a descoperit recent câteva sute de gene care au fost implicate în tulburarea de spectru autist (pentru detalii, a se vedea Krishnan et al., 2016). Alte echipe de cercetare utilizează algoritmi similari pentru a identifica contribuțiile genetice la boala Alzheimer, afecțiunile renale și cancer.

Aceste tehnici de editare a genelor ar putea fi utilizate și pentru a testa propunerea lui Sacks și a colegilor săi, precum și sugestiile lui Bray și ale mele. Identificarea genelor care controlează crearea membranelor celulare sau a celor care codifică mecanismul de monitorizare a temperaturii ar contribui mult la înțelegerea modului în care aceste funcții au loc și cum sunt integrate în viața fenomenală a organismului. Cercetătorii ar ști când au găsit genele sau secvența (secvențele) de ADN potrivite, deoarece funcțiile coordonate ar fi modificate sau ar înceta după ștergere. Genomul mai multor tulpini de *E. coli* (și al altor bacterii) a fost secvențiat și sunt cunoscute funcțiile multora dintre cele 4 000-5 500 de gene (fiecare tulpină are un genom distinct). Până în prezent, nimeni nu le-a căutat pe cele care codifică pentru proteinele responsabile de tipurile de funcții despre care am discutat. Dar, în principiu, nu există motive pentru care un astfel de proiect să nu poată fi realizat.

Am subliniat mai devreme că o piatră de poticnire pentru toate teoriile privind originea conștiinței, inclusiv modelul **CBC**, este *dilema emergentistului*. Speculațiile ghidate biologic ale lui Bray, Sacks și ale colegilor, precum și cea pe care tocmai am oferit-o, sunt exemple de unde cred că se va găsi rezolvarea. Este posibil, de fapt cred că este probabil, ca viața și sensibilitatea să fi coexistat atunci când supa prebiotică a fost transformată în supă abiotică în urmă cu aproximativ 3,7 miliarde de ani. Viitorul ar trebui să fie interesant, pe măsură ce cercetătorii continuă să facă progrese în problema originii vieții, biochimiștii perfecționează tehnicile CRISPR și sunt dezvoltați algoritmi AI mai puternici pentru analizarea seturilor mari de date create de genetica modernă.

---

77 Tehnica CRISPR este relativ nouă, dar, împreună cu procedurile de editare a perechilor de baze, care sunt și mai eficiente pentru a viza gene unice, au demonstrat deja o capacitate extraordinară de a modifica genomul organismelor individuale. Aceste tehnici ridică probleme bioetice profunde. Agențiile guvernamentale din țările care sprijină cercetarea genetică de vârf și organismele științifice din întreaga lume se străduiesc să le abordeze și să ajungă la orientări care să fie funcționale, corecte și etice. Noua carte a lui Jennifer Doudna și Samuel Sternberg, *A Crack in Creation* (O fisură în creație), este cel mai recent efort de a aborda o tehnică care, după cum spune Doudna, face genele unui organism „aproape la fel de editabile ca o simplă bucată de text”.

Cum dezvoltă organismele pluricelulare o conştiinţă unitară?

Aceasta este o întrebare, dar există de fapt două probleme care trebuie abordate. Prima este cea mai elementară întrebare: Cum au evoluat organismele pluricelulare? Aceasta nu este nouă; a fost mult timp urmărită de microbiologi şi, de fapt, este considerată a fi una dintre principalele probleme doar parţial rezolvate în biologia evolutivă (a se vedea Pennisi, 2018). A doua întrebare, cum să oferim unui organism pluricelular o experienţă fenomenologică singulară, nu a fost, din câte mi-am putut da seama, pusă în mod serios. Această lipsă de preocupare nu este teribil de surprinzătoare, deoarece, dacă nu îţi stabileşti tipul de poziţii pe care le are cadrul **CBC**, este puţin probabil să apară. Trebuie să porniţi de la presupunerea că organismele unicelulare au o conştiinţă rudimentară pentru a vă întreba cum se integrează numeroasele experienţe conştiente individuale ale fiecărei celule dintr-un organism pluricelular într-o experienţă singulară a întregului. Să începem cu o privire asupra primei întrebări, deoarece este cea la care există o literatură extinsă.

Cea mai probabilă secvenţă de evenimente consideră că primele organisme pluricelulare s-au dezvoltat din colectivele formate de multe specii unicelulare şi, ceea ce este important, au rămas astfel. Această coeziune este esenţială deoarece, dacă colectivele nu au succes şi celulele individuale nu sunt pe deplin integrate, devine prea uşor pentru ele să se retragă şi să se întoarcă la o viaţă independentă. Mai multe tulpini de drojdie sunt suspectate că au urmat această cale şi au revenit la viaţa de organisme individuale. Trebuie să existe un avantaj adaptiv care provine din rămânerea cu entitatea multicomponentă sau un mecanism sau mecanisme care fac dificilă revenirea la viaţa ca organisme separate.

Mai multe tehnici de realizare a acestui lucru sunt cunoscute ca fiind utilizate şi au fost observate la speciile unicelulare contemporane. Într-o procedură, o celulă se divide, iar celula fiică, în loc să înoate, se ataşează de celula mamă. O serie de repetări ale acestei lipiri şi începe să se formeze o colonie. Într-o altă procedură, un grup de celule separate se adună şi secretă substanţe lipicioase care le ţin împreună - aşa cum am văzut că se întâmplă mai devreme când am analizat modul în care unele specii de bacterii formează covoare microbiene şi biofilme. Într-un alt caz, după divizarea unui nucleu, se formează o membrană separată în jurul celui de-al doilea pachet de ADN, rezultând două celule distincte în cadrul unui singur perete celular. Mai multe iteraţii ale acestui proces şi suntem pe drumul cel bun. Dar faptul că toate aceste procese sunt cunoscute nu este suficient. În fiecare caz este nevoie de ceva mai mult. Trebuie să existe procese biologice care să ofere colectivităţii un avantaj adaptiv şi altele care să reducă probabilitatea ca indivizii să se defecteze.

Un mecanism care a fost propus se bazează pe principiile simbiotice în care celulele individuale dobândesc funcţii care sunt utile colectivităţii. După cum a susţinut Margulis, este probabil ca acest tip de cooperare să fi apărut atunci când două sau mai multe specii unicelulare au devenit dependente una de cealaltă - poate pentru că fiecare secreta o anumită moleculă care aducea un beneficiu celeilalte. În timp, este probabil ca aceste specii să fi dezvoltat această dependenţă reciprocă până în punctul în care nu pot supravieţui cu uşurinţă una fără cealaltă şi, odată cu repetarea acestei dependenţe moleculare, colonia în creştere ajunge să formeze un colectiv pluricelular. Un astfel de proces îndeplineşte cerinţele unui organism pluricelular emergent şi robust. Colectivul este adaptiv; rata sa de supravieţuire va fi mai mare decât a oricărei specii izolate; iar defectarea celulară va fi redusă, având în vedere că părăsirea colectivului înseamnă renunţarea la câştigurile care provin din frecventarea tuturor celorlalte celule utile şi a secreţiilor lor utile.

Acest proces este un exemplu a ceea ce biologii numesc un *clichet*, deoarece, după un astfel de progres, devine extrem de dificil, din punct de vedere evolutiv, să dai înapoi. Prin urmare, probabilitatea ca acest nou colectiv pluricelular să supravieţuiască este crescută şi, după cum a susţinut Margulis, este mai probabil ca, în cele din urmă, să devină o specie distinctă.[78] Acest tip de

78 Margulis, după cum am menţionat în capitolul 2, a fost unul dintre adevăraţii iconoclaşti ai domeniului. Ea a considerat că aceste tipuri de relaţii simbiotice, deoarece se întâlnesc între specii şi chiar între filii, reprezintă forţa

proces simbiotic ar putea avea loc şi în cadrul unei singure specii dacă două sau mai multe mutaţii fortuite ar crea celule specializate care ar putea secreta molecule care ar oferi beneficii reciproce. Un colectiv format din aceste celule cooperante, indiferent de numărul de specii individuale implicate, ar fi pe cale să devină o nouă specie pluricelulară.

Trebuie să reţinem că procesul de creare a organismelor multicelulare viabile a durat foarte mult timp. După cum am menţionat anterior, primele dovezi de viaţă pe planetă datează de acum 3,7 miliarde de ani. Primele dovezi fosile ale formelor pluricelulare au o vechime de aproximativ 600 de milioane de ani - deşi este probabil ca unele specii pluricelulare să fi evoluat mai devreme, dar, deoarece nu aveau părţi rigide ale corpului, este puţin probabil să fie păstrate în arhivele fosile. De asemenea, este posibil ca, în acest interval de timp, unele specii pluricelulare timpurii să nu se fi adaptat la capriciile mediului şi să fi renunţat pur şi simplu şi să fi revenit la forma lor iniţială unicelulară. Dar, indiferent de modul în care se determină punctul de origine, este absolut clar că crearea unei specii multicelulare viabile a fost o problemă destul de dificilă de rezolvat pentru acele versiuni unicelulare iniţiale şi a fost nevoie de miliarde de ani pentru a face acest lucru.

Această problemă, oricât de interesantă ar fi, nu este cea centrală pentru meditaţiile noastre de aici. Din punctul de vedere al **CBC**, problema profundă nu este determinarea mecanismelor de formare a speciilor multicelulare viabile. Este vorba despre modul în care funcţiile experienţiale, sensibilitatea care era operaţională în fiecare celulă individuală, a ajuns să fie cooptată de colectiv. Este vorba despre înţelegerea modului în care întregul organism pluricelular simte, de exemplu, durerea ca pe o experienţă integrată, coerentă şi aversivă şi nu ca pe un sentiment de tip pointillist al unui număr mare de experienţe individuale ale durerii. Este modul în care entitatea pluricelulară ajunge să decidă să meargă înainte pentru că acolo sunt detectate niveluri mai ridicate de nutriţie, în loc ca fiecare celulă să ia propriile decizii cu privire la locomoţie şi mişcare.

Am dat un indiciu mai devreme, în capitolul 4, care se referă la această problemă. Atunci când am analizat comunicarea în cadrul colectivităţilor de specii unicelulare şi între acestea, am văzut care este un mecanism probabil pentru această schimbare epistemică. Amintiţi-vă de studiul realizat de Gürol Süel şi colaboratorii săi de la Universitatea din California San Diego, în care au descoperit că o colonie a bacteriei *Bacillus subtilis* poate, prin secreţia sincronă de ioni încărcaţi, să comunice unui alt colectiv informaţii despre starea sa nutriţională generală. Atunci când acest lucru se întâmplă, dintr-o perspectivă funcţională, colectivul „vorbeşte”, ca să spunem aşa, pentru toate celulele individuale care îl alcătuiesc. Celulele din colonia care transmite mesajul se confruntă cu deficite nutritive şi, împreună, îşi coordonează şi îşi transmit starea. A doua colonie primeşte mesajul şi, funcţionând ca un fel de *gestalt* slab format, îşi reduce propriile rate de reproducere pentru a răspunde nevoilor primei.[79]

Ceea ce se întâmplă cu bacteria lui Süel este o acţiune coordonată a unui colectiv, luată în numele fiecăruia dintre organismele individuale care îl compun. O astfel de abilitate este în mod evident adaptivă şi, în timp, ar deveni fixată în ADN-ul speciei. Acest tip de strategie *„lasă grupul să vorbească în numele indivizilor*”, odată stabilită, ar deveni fundamentul pentru ca celelalte funcţii de bază să fie transferate, ca să spunem aşa, de la indivizi la grup. Odată ce specia a adoptat acest mecanism, următorul pas evident ar fi ca anumite celule din colectiv să dezvolte trăsături specializate care ar putea fi utilizate în beneficiul tuturor componentelor celulare individuale. Un grup de celule ar putea dezvolta un modul mai eficient de detectare a nutrienţilor şi ar putea utiliza un mecanism de secreţie a moleculelor pentru a comunica celorlalte celule din colonie informaţii despre locul în care trebuie să se deplaseze pentru a găsi hrană. Restul coloniei ar avea de câştigat,

---

motrice principală în evoluţie, chiar mai mult decât selecţia naturală darwinistă. Majoritatea biologilor evolutivi obişnuiţi recunosc existenţa mecanismelor simbiotice în evoluţie, dar nu sunt de acord că acestea au avut un impact mai mare asupra speciei decât selecţia naturală.

79 Interesant este faptul că acest tip de cooperare are analogii funcţionale cu tipul Margulis de relaţie simbiotică. Ambele proceduri sugerează apariţia unor tipuri foarte elementare de cooperare şi altruism, aspecte la care vom reveni mai târziu în acest capitol.

deoarece nu ar mai trebui să cheltuiască propriile resurse metabolice pentru detectarea gradienților de concentrație a zahărului, ci pur și simplu să preia informațiile necesare din mesajele trimise de celulele care au devenit, de fapt, monitorul nutrienților colectivului.

Se poate scrie o poveste similară pentru alte funcții senzoriale, motorii și emoționale. Celulele care, printr-o variație naturală sau o mutație fortuită, se întâmplă să dezvolte un avantaj în rolurile perceptive, poate prin faptul că au mai multe molecule de rodopsină, ar deveni primele sisteme de detectare a luminii. Celelalte celule din colectiv își pot abandona acum funcțiile foto-senzoriale costisitoare din punct de vedere metabolic și își pot folosi resursele biochimice pentru a se specializa în alte operațiuni, poate termo-senzori sau locomoție. Mutațiile ar putea produce celule cu sensibilitate crescută la stimuli dureroși și nocivi și ar putea evolua în *nociceptori*. Nu este greu de imaginat cum acest mecanism, odată stabilit, ar putea genera o specie complexă, pluricelulară, cu unități celulare specializate pentru anumite funcții biologice - chiar dacă ar dura trei miliarde de ani.

Desigur, aceasta este o speculație, dar se bazează pe principii bine stabilite ale biologiei celulare și pe cercetări recente privind modul în care specii atât de primitive precum *Bacillus subtilis* pot comunica la distanță. De asemenea, ea dezvăluie paralele evidente cu mecanismele oferite de Bray pentru modul în care celulele individuale au ajuns să dezvolte sisteme senzoriale bazate pe cunoaștere. O prezint aici ca o schemă ipotetică pentru modul în care funcțiile experiențiale ale organismelor unicelulare din cadrul unui colectiv ar putea fi transformate, astfel încât colonia să aibă o experiență unitară, mai degrabă decât o serie de reacții senzitive separate, necoordonate, din partea fiecărei componente celulare.

Dar să continuăm să aprofundăm acest cadru pentru a vedea cum a fost probabil instanțiat și cum ar putea funcționa în specii mult mai complexe, precum noi. Din perspectiva abordării **CBC**, se presupune că celulele individuale ale corpului nostru au niveluri de sensibilitate asemănătoare bacteriilor. Să presupunem că se întâmplă ceva care provoacă leziuni ale țesuturilor celulelor musculare din piciorul dumneavoastră. Printre celulele din zonă se numără unele care au dezvoltat *nocicepția* - un sentiment subiectiv aversiv, la nivel celular, că ceva nu este în regulă. Cu toate acestea, deoarece toate fac parte dintr-un colectiv mare (dumneavoastră), structura lor genetică le determină să comunice experiența lor altor celule, în special neuronilor aferenți locali, în mod coordonat, astfel încât să semnaleze experiența lor de durere.

Odată ce suficient de multe dintre ele au făcut acest lucru, nervii aferenți se declanșează și mesajul ajunge în locul în care ajung toate aceste tipuri de lanțuri de informații, creierul tău. Colectivul tău are acum o experiență singulară a durerii, cu informații relevante despre unde și cât de mult. Ceea ce devine parte a conștiinței voastre este o experiență integrată. Reacțiile separate ale fiecărei celule componente, oricât de reale ar fi ele, nu fac parte din simțul vostru fenomenal.

Un alt exemplu al acestei legături între sistemele mai vechi, evolutive, de nivel inferior și conștiință este modul în care funcționează sistemul implicit la om. Aici, informațiile sunt preluate (învățare implicită) și stocate (memorie implicită) fără conștientizarea proceselor de achiziție sau a existenței amintirilor. Aceste funcții au fost atribuite unor structuri cerebrale mai vechi din punct de vedere evolutiv (pentru detalii, a se vedea R T. Reber, 2013) care „își fac treaba” în fundal. Aceste structuri cerebrale și-au păstrat funcțiile de procesare, dar și-au predat, ca să spunem așa, rutele către conștiință către alte zone, cum ar fi cortexul prefrontal. Modelul de cunoaștere implicită dezvoltat în altă parte (Reber, 1993) subliniază modul în care principiile biologiei evolutive operează în aceste mecanisme.

Aceste secvențe de comunicare de la celulă la celulă sunt extensii evolutive directe ale situației despre care presupunem că a avut loc în urmă cu o jumătate de miliard de ani, când colectivitățile pluricelulare au devenit organisme distincte și au descoperit modalități de a coordona și combina multitudinea de experiențe ale componentelor lor celulare în experiențe subiective unitare. Un motiv pentru care este probabil să fi avut loc acest tip de procese evolutive este acela că funcțiile fiziologice sunt costisitoare din punct de vedere metabolic. Acestea utilizează resurse, iar

atunci când costul acestora depăşeşte funcţiile adaptative pe care le susţin, ele devin fundături darwiniene. O soluţie este specializarea şi utilizarea formelor şi funcţiilor evoluate anterior ca platforme pentru cele noi care au un avantaj adaptiv în procesele continue de selecţie naturală.

La fel ca în cazul întrebării anterioare privind modul în care a apărut prima dată sensibilitatea primordială, nu există răspunsuri definitive în acest caz. Dar cadrul speculativ prezentat aici este o consecinţă clară a modelului **CBC** şi poate fi testat empiric. Ştim unele lucruri despre modul în care au evoluat speciile multicelulare şi, poate, în această poveste biofizică va fi scrisă într-o zi cea bioexperienţială.

Cum rămâne cu plantele?

Trebuie să luăm în considerare flora împreună cu fauna. Am abordat această problemă mai devreme, dar ea necesită mai multă dezvoltare şi discuţii. Una dintre consecinţele evidente ale poziţiei **CBC** este că întrebarea dacă plantele sunt conştiente este inevitabilă. Majoritatea numeroşilor alţi filosofi şi oameni de ştiinţă care au abordat tipurile de întrebări şi probleme pe care le-am discutat în această carte pot (şi o fac) pur şi simplu să ocolească problema. Deoarece punctul de plecare al celor mai mulţi neuroştiinţifici şi filosofi ai minţii este conştiinţa umană, aceasta nu apare niciodată. Ceilalţi, cei care caută punctul din marele arbore al vieţii în care conştiinţa ajunge pentru prima dată pe scenă, se uită aproape invariabil în sus şi în jos în regnul animal şi nu iau niciodată în considerare acele fotosintetizatoare cu pereţi celulari rigizi ca subiecte de explorare. Biofilosofii pur şi simplu îşi expun punctele de vedere, îşi prezintă argumentele şi îşi resping sau încearcă să îşi respingă adversarii. Cercetătorii mai orientaţi empiric îşi desfăşoară studiile, examinează orice cohortă sau specie sau sindrom de interes, îşi colectează datele şi ne spun ce au descoperit. Dacă nu menţionaţi plantele, este destul de puţin probabil ca cineva să semnaleze acest lucru, dacă nu din alt motiv decât că nici ei nu vor să vorbească despre ele.

Dar trebuie să o facem aici. După cum s-a menţionat în capitolul anterior, din cauza cronologiei evoluţiei şi a naturii argumentelor de bază care stau la baza teoriei **CBC**, ar fi pur şi simplu necinstit din punct de vedere intelectual să încercăm această manevră. Teoria **CBC** plasează originile sensibilităţii, ale minţii, chiar la început, odată cu apariţia vieţii în sine, în organismele unicelulare şi în cele mai primitive dintre acestea, procariotele şi *arhaea*, unde nucleul încapsulat nu a evoluat încă. Plantele sunt eucariote şi sunt pluricelulare. Ele au evoluat cu mult după momentele evoluţiei asupra cărora ne-am concentrat. Şi, deoarece am folosit principiul evoluţionist al continuităţii funcţiei în mai multe locuri, ne confruntăm cu o enigmă. Conform cronologiei şi invocând acest principiu, plantele ar trebui să fie conştiente, să aibă stări mentale, sensibilitate. Dar eu nu sunt sigur că acest lucru este corect. Şi trebuie să încerc să vă conving de ce - şi, desigur, să mă conving şi pe mine în acelaşi timp. Cred că există trei elemente cheie aici:

- *motilitate*
- *structura peretelui celular*
- *costul metabolic al sensibilităţii*

Să analizăm modul în care toate aceste trei elemente şi-au putut juca rolul în procesul prin care plantele au evoluat şi au făcut-o fără a fi sensibile.

<u>Mobilitatea</u>.

Speciile unicelulare cu mobilitate se deplasează în lume şi nu este vorba doar de o rătăcire fără ţintă. După cum s-a menţionat anterior, aceasta este caracterizată cel mai bine ca o deplasare aleatorie părtinitoare care este determinată de evaluarea condiţiilor interne în comparaţie cu

elementele externe. Speciile mobile vor trebui să interacţioneze cu mediul, nu doar să acţioneze în cadrul acestuia. După cum am subliniat în capitolul 4, atât sensibilitatea la condiţiile interne ale unui organism unicelular, cât şi cunoştinţele despre condiţiile externe din lumea înconjurătoare sunt esenţiale în acest caz - şi ambele trebuie să fie deţinute şi comparate între ele.

Dar dacă un organism nu se mişcă, dacă mediul în care există este ceea ce este şi sarcina sa este să facă tot ce poate în aceste condiţii, nu există niciun beneficiu de pe urma acestor capacităţi senzoriale. Plantele nu au mişcări voluntare dirijate. Ele pot fi supuse relocării de către maree, curenţi, vânturi sau animale, dar sunt călătoare pasive, care au puţin sau deloc de spus în ceea ce priveşte locul în care ajung. Mai mulţi teoreticieni pe care i-am analizat mai devreme, precum Maxine Sheets-Johnstone şi Antonio Damasio, au identificat mişcarea întrupată ca fiind o caracteristică esenţială a subiectivităţii, o funcţie critică care stă la baza conştientizării şi conştiinţei. Şi, aşa cum au subliniat Sacks, Damasio, Cook, Carvalho şi Hunt în scrisoarea lor către Koch,

La protozoarele unicelulare, majoritatea răspunsurilor sunt mişcări simple ca răspuns la un stimul chimic în flageli, cili sau pseudopode, care sunt determinate de concentraţiile de ioni din citoplasmă. La animalele care conţin multe miliarde de celule receptoare senzoriale, neuroni şi celule musculare, răspunsurile sunt coordonate în activitatea motorie orientată spre un scop a anumitor musculaturi.

În organismul unicelular, mişcarea dirijată este legată de experienţă şi nu poate fi separată de aceasta. Fără motilitate nu ar exista niciun beneficiu adaptiv din stările endogene, subiective. Ce rost are să ştii unde sunt nutrienţii dacă nu poţi ajunge acolo? Ce câştig ar avea detectarea gradienţilor de temperatură dacă eşti blocat în starea termică momentană din mediul în care ţi-ai pus rădăcinile? Care este avantajul evolutiv de a fi conştient dacă eşti incapabil de mişcare dirijată?

Botaniştii observă că există o formă de mişcare la plante, cum ar fi modul dirijat în care sistemele de rădăcini cresc ca răspuns la nivelurile de nutrienţi sau modul în care speciile asemănătoare viţei de vie extind vreascuri care răspund la obiectele pe care le ating. După cum vom vedea mai jos, unii botanişti precum Anthony Trewavas de la Universitatea din Edinburgh susţin că, deşi aceste mişcări diferă de cele ale animalelor prin faptul că sunt foarte lente, ele se califică drept motilitate. Trewavas este, de asemenea, de părere că plantele sunt, de fapt, sensibile. Vom analiza mai detaliat poziţia sa mai jos. Vom vedea, de asemenea, că aceasta este împărtăşită de mai mulţi biologi ai celulelor vegetale.

Structura peretelui celular.

Plantele, conform modelului endosimbiotic general acceptat în prezent, au evoluat în urma unui act de încorporare în urmă cu aproximativ 1,6 miliarde de ani, când o cianobacterie verde (o specie procariotă care dezvoltase fotosinteza pentru extragerea energiei din lumina soarelui) a fost încorporată de o celulă eucariotă. Rezultatul a fost celula alga verde şi începutul regnului *Plantae*. Analizele genetice efectuate de botaniştii Dana Price, Debashish Bhattacharya şi colegii lor de la Universitatea Rutgers sugerează, poate surprinzător, că acest lucru s-a întâmplat o singură dată, iar toate speciile de plante existente pot fi înţelese ca evoluând din această unică himeră.

Problema de interes aici este faptul că această nouă specie a ajuns să aibă peretele celular rigid al „părintelui” procariot. După cum au remarcat Iben Serensen şi colegii de la Universitatea din Copenhaga, deşi evoluţia diversităţii enorme a pereţilor celulari ai plantelor nu este încă pe deplin înţeleasă, rigiditatea acestora este o caracteristică definitorie a tuturor speciilor de plante. Pereţii celulari rigizi funcţionează în mai multe moduri, inclusiv prin creşterea rezistenţei la tracţiune pentru a susţine structura plantei, protecţia împotriva bolilor, furnizarea unei structuri poroase pentru circulaţia apei şi a nutrienţilor şi permiterea stocării moleculelor pentru creştere şi comunicare între părţile organismului. Acestea sunt, de asemenea, legate de lipsa de motilitate şi, după cum s-a menţionat, adoptarea unei existenţe sesile face mai puţin probabil ca astfel de specii să continue să îşi menţină subiectivitatea internă, primordială.

Cu toate acestea, după cum mi-a semnalat Frantisek Baluska, de la Universitatea Bonn din Germania, într-un schimb de e-mail, pereții celulari rigizi ai plantelor pot fi responsabili pentru lipsa „aparentă" de motilitate, dar povestea este mai complicată. De exemplu, modul în care cresc sistemele radiculare poate fi văzut ca un analog al funcțiilor de mișcare dirijată ale speciilor animale. Motilitatea organismelor unicelulare, a observat el, face ușor de observat modul în care acestea își ajustează căile pentru a găsi climate mai confortabile, medii mai nutritive și medii mai puțin nocive. Plantele fac lucruri similare, dar foarte lent. Sistemul radicular, pe care Baluska îl aseamănă cu creierul plantei, atinge obiective similare prin evaluarea temperaturii, a nivelului de nutrienți și a existenței substanțelor toxice și ajustează modelele de creștere într-o direcție mai degrabă decât în alta. La fel ca Trewavas, el consideră că flora este sensibilă. Din nou, mai multe despre acest lucru mai jos.

Costul sensibilității.

Dintr-o perspectivă evolutivă, ar fi aproape sigur adaptabil ca aceste prime plante să facă cea mai rară dintre mișcările biologice: să piardă o funcție cândva utilă și să renunțe la capacitatea primitivă de conștientizare. Sentimentul este costisitor din punct de vedere metabolic. Ai nevoie de receptori specializați, de mecanisme interne de procesare a informațiilor, de un aparat de control motor pentru ghidarea mișcării și de capacitatea de a monitoriza atât stările interne, cât și mediul înconjurător și de a coordona totul eficient. O specie căreia îi lipsește motilitatea și-ar spori avantajul evolutiv prin canalizarea cheltuielilor metabolice de la unele dintre aceste funcții către alte capacități. Pe scurt, este posibil ca plantele să fi fost singurul eveniment evolutiv major în care primele specii nemotile au abandonat *sensibilitatea* primordial al predecesorilor lor și și-au redirecționat resursele metabolice în alte moduri utile.

Pe de altă parte.

Acum, după ce m-am angajat în ceea ce cred că este un argument rezonabil de bine construit, dar în mod clar speculativ, care conduce la sugestia puternică, dacă nu la concluzia, că plantele nu sunt conștiente, este corect să aducem în discuție munca celor care au o abordare foarte diferită, care conduce la un punct final foarte diferit. Un număr de botaniști proeminenți sunt activi în promovarea talentelor variate ale plantelor și în prezentarea lor într-o manieră care pare să reflecte multe dintre funcțiile pe care le-am identificat în capitolul 4 ca fiind cerințe pentru sensibilitate.

Naturalistul britanic Richard Mabey și Michael Pollan, profesor de jurnalism la Universitatea Berkeley din California, au scris cărți de succes pe această temă, iar Robert Lanza, un cercetător medical cunoscut mai ales pentru activitatea sa inovatoare privind celulele *stem* și clonarea, a prezentat un rezumat al abilităților plantelor care au fost identificate (până în prezent):

Plantele au, de asemenea, rețele neuronale rudimentare și capacitatea de percepție primară. Într-adevăr, planta Drosera (Drosera) va prinde o muscă cu o precizie incredibilă - mult mai bine decât o poți face cu un strivitor de muște. Unele plante știu chiar și când furnicile se apropie de ele pentru a le fura nectarul și dispun de mecanisme pentru a se închide atunci când acestea se apropie. Oamenii de știință de la Universitatea Cornell au descoperit că, atunci când un vierme începe să mănânce Artemisia tridentata (Sagebrush), planta rănită emite un miros care avertizează plantele din jur ... că problemele sunt pe drum. Aceste plante, la rândul lor, pregătesc apărarea chimică care trimite creaturile înfometate în direcția opusă.

Biologul celular Baluska împreună cu fiziologul vegetal Stefano Mancuso de la Universitatea din Florența și colegii lor au explorat ceea ce ei consideră a fi diversele abilități

neapreciate ale plantelor. În ceea ce bănuiesc că este un pic de spirit ludic, ei și-au numit abordarea „neurobiologie vegetală” și se concentrează pe paralelele dintre diferite caracteristici ale plantelor și componentele moleculare care stau la baza mecanismelor neuronale la animale. Ei observă, de exemplu, că multe dintre proteinele responsabile de propagarea impulsurilor neuronale la animale se găsesc în plante și, în mod interesant, în procariote. Baluska a fondat două reviste concepute ca piețe de desfacere pentru cercetarea acestor tipuri de probleme, inclusiv o lucrare recentă din laboratorul lor care sugerează că unele specii de plante pot vedea de fapt, că celulele epidermice superioare ale frunzelor pot acționa ca o lentilă care focalizează lumina asupra celulelor subepidermice.[80] În 2011, Trewavas și Baluska au prezentat o vedere de ansamblu a elementelor biofiziologice și comportamentale comune dintre plante și animale. Concluzia lor:

Comportamentul plantelor este activ, determinat de un scop și intenționat. Prin capacitatea sa de autorecunoaștere și de rezolvare a problemelor, la fel ca celelalte organisme descrise în acest articol, este, prin urmare, adaptiv, inteligent și cognitiv.

Un alt cercetător activ în acest domeniu, Monica Gagliano de la University of Western Australia, își caracterizează în mod specific abordarea ca fiind una care extinde, în cuvintele sale, „*conceptul de cogniție la plante, relansând discursul privind subiectivitatea, sensibilitatea și poziția etică a plantelor*”. În mai multe studii recente, ea și colaboratorii săi au prezentat dovezi care sugerează că plantele sunt capabile de învățare asociativă, au memorie, sunt sensibile la sunete și sunt comunicatori eficienți. Biochimistul Michael Gross a sintetizat situația:

În mod tradițional, vegetația este considerată pasivă, nefăcând nimic altceva decât ceea ce este esențial pentru a crește și a supraviețui. Cu toate acestea, se acumulează dovezi în sprijinul unor noțiuni altădată ezoterice, conform cărora plantele pot comunica, își pot aminti, chiar și număra - caracteristici pe care le-am numi cognitive dacă ar fi observate la animale.

Plantele au dezvoltat, de asemenea, abilități farmaceutice remarcabile și produc o serie de compuși pentru protecție, pentru a semnala pericolul, pentru a otrăvi prădătorii, pentru a atrage păsările și insectele, pentru a facilita fertilizarea și pentru a încuraja animalele să transporte semințe pentru propagare la distanță. Înclinația lor de a dezvolta acești compuși diverși este susținută de unii ca un indicator al faptului că plantele au experiențe interne și au dezvoltat aceste medicamente pentru a evita să fie puse în circumstanțe care sunt subiectiv neplăcute. Aceste talente sunt, desigur, exact tipurile de funcții la care ne-am aștepta să fi evoluat la speciile sesile.

Ce să facem cu aceste constatări și cu interpretările entuziaste ale multora? Se recomandă prudență. Merită să ne amintim de o perioadă anterioară de afirmații entuziaste despre plante, exemplificată de cartea lui Peter Tompkins și Christopher Bird din 1973, *The Secret Life of Plants* (Viața secretă a plantelor). Acest volum a contribuit foarte puțin la consolidarea cercetării științifice și a provocat multe daune biologilor și botaniștilor serioși care lucrau pentru a desluși complexitatea regnului vegetal. Volumul lui Tompkins și Bird descria experimente care ar fi demonstrat că plantele au emoții, pot detecta sentimentele altor plante și chiar ale oamenilor și pot identifica vinovăția observatorilor care au comis acte de violență (cum ar fi distrugerea unei alte plante). *The Secret Life of Plants* a devenit o carte de succes, a fost ecranizată și, bineînțeles, s-a dovedit curând că se baza pe o știință total falsă. O examinare mai atentă a metodologiilor utilizate în cadrul studiilor a arătat că acestea erau fatalmente eronate și că niciunul dintre rezultatele pretinse nu putea fi reprodus. Michael Gross și-a încheiat prezentarea din 2016 cu aceste cuvinte de precauție cu privire la referirile îndrăznețe la cunoașterea plantelor:

---

80 În mod interesant, această posibilitate a fost luată în considerare și de Francis Darwin, fiul lui Charles. În cartea lor scrisă în colaborare în 1880, *The Power of Movement in Plants* (Puterea mișcării la plante), aceștia reflectează asupra unei mari varietăți de posibile caracteristici sensibile ale plantelor.

Riscul unui astfel de limbaj este că poate reînvia controversele legate de afirmațiile extrem de ezoterice făcute în trecut cu privire la conștiința plantelor, adesea bazate pe un amestec de cunoștințe științifice și pură fantezie. Descoperirile făcute recent cu ajutorul celor mai avansate tehnici pe care știința secolului XXI le poate pune la dispoziție sugerează, totuși, că abilitățile cognitive [ale plantelor] merită să fie luate în serios și studiate în continuare.

Este demn de remarcat faptul că, cu excepția lucrării acum demascate a lui Tompkins și Bird, cei care au adoptat poziția conform căreia plantele sunt sensibile se concentrează în primul rând pe elemente specifice ale sensibilității, în special pe cogniție. Am observat această prejudecată mai devreme, când am analizat teoretizarea despre sensibilitate în cadrul funcționaliștilor independenți de hardware în *capitolul 1* și am văzut-o din nou în eforturile de a identifica momentul în care a apărut conștiința în evoluție în *capitolul 3*. Practic, toate cercetările și speculațiile s-au axat pe funcții cognitive precum percepția, învățarea, memoria și luarea deciziilor, neglijând afectele, emoțiile și sentimentele. Majoritatea cercetărilor efectuate de acești neurobiologi moderni ai plantelor au o viziune similară.

Cu siguranță nu a fost întotdeauna așa. În legende și mituri, plantele sunt văzute ca întruchipări ale afectului, inserându-se în viața oamenilor, exprimând emoții, creând rivalități și jucând adesea roluri semnificative în miturile creației. În remarcabila carte a lui Richard Folkard[81] din 1884, Plant Lore, *Legends, and Lyrics: Embracing the Myths, Traditions, Superstitions, and Folk-lore of the Plant Kingdom*, plantele și copacii erau, în cuvintele sale,

... considerate aproape universal fie ca ființe simțitoare, fie ca locuințe ale unor spirite a căror existență era legată de viața copacilor pe care îi locuiau. De aici au apărut concepțiile despre ... spirite binefăcătoare care populau pădurile și locuiau în copaci individuali ... unde, precum Puck, [se] jucau la lumina lunii în pădure.

Această noțiune a plantelor care prezintă acest tip de trăsături, pe care le asociem în mod normal cu oamenii, a fost o temă dominantă în societățile dinaintea iluminismului. Așa cum Folkard analizează cu drag literatura antropologică, flora pătrundea în viața oamenilor, provocând durere, vindecând boli, aranjând legături, prevenind (și instigând) războaie și conflicte, transformând în mod magic frunzele în pietre, schimbându-și forma pentru a încurca planurile celor puternici. Dar, ceea ce este fascinant, practic niciodată nu au fost descriși ca maeștri ai numerelor, rezolvatori raționali de probleme, jonglatori atenți cu alternativele înainte de a lua decizii sau exploatând logica pentru a câștiga un argument.

Cred că ceea ce vedem în activitatea botaniștilor contemporani interesați de posibilitatea sensibilității plantelor face parte din această tendință de a vedea funcțiile minții și conștiinței ca fiind reflectarea proceselor cognitive. După cum am observat atunci când am analizat eforturile comunității IA și ale filosofilor care au îmbrățișat versiunea independentă de hardware a funcționalismului, prezumția a fost că, dacă un artefact efectuează calcule identice cu cele ale unui creier uman, atunci acesta ar manifesta conștiință adevărată. Cu excepția câtorva specialiști în neuroștiințe, precum Antonio Damasio, se vorbea puțin sau deloc despre afect, emoție, sentimente. Într-o recenzie recentă, Stanislaw Daheane și colegii săi analizează, încă o dată, subiectul mai larg al conștiinței și pun vechea întrebare: „... ar putea mașinile să o aibă?” Acoperirea lor a domeniului

81 Magnum opus-ul lui Folkard este un compendiu de 650 de pagini al tuturor miturilor sau legendelor culturale cunoscute, care se referă la modul în care plantele își exprimă emoțiile și sentimentele și se implică în viața oamenilor care trăiesc printre ele. Nu am putut găsi nicio dovadă că ar fi avut vreo pregătire în botanică, literatură sau mitologie. Într-o sursă a fost identificat ca „tipograf” și, după cum notează în prefață, el însuși a produs volumul. Acesta conține multe greșeli de ortografie, nenumărate greșeli de scriere și mai mult decât câteva fraze indescifrabile. Cu toate acestea, este o plăcere absolută să o parcurgi.

este extinsă și include peste 90 de referințe. Nu apar nicăieri cuvintele „afect” sau „emoție”. Sunt luate în considerare doar operațiunile computaționale.

Acest accent îngust este de înțeles; funcțiile cognitive implică calcule. Artefacte precum computerele digitale puternice sunt foarte bune la acestea. Emoția, motivația, afecțiunea, suferința, bucuria nu se pretează cu ușurință la aceste tipuri de reprezentări digitale. Un AI poate fi capabil să simuleze acțiunile unui matematician care a descifrat o teoremă nerezolvată anterior, dar nu poate emula bucuria descoperirii.

Dar să revenim la adevărata problemă. Sunt plantele sensibile? Poziția mea este una de scepticism studiat. După cum am menționat mai sus, cred că motilitatea autodirijată a întregului organism este cu siguranță un factor important, la fel ca și problema costului metabolic al sensibilității. Și, cu o notă de ironie, voi împrumuta linia de argumentare prezentată de Dan Dennett împotriva principiilor de bază ale modelului meu **CBC**. După cum am menționat anterior, în cartea *From Bacteria to Bach and Back*, Dennett a susținut că toate talentele aparent mentale ale speciilor unicelulare precum bacteriile și alte procariote sunt simple reacții mecanice care rezultă din predispoziții genetice apărute cu eoni în urmă, pe măsură ce aceste creaturi primitive evoluau. El susține că nu există nicio dovadă de sensibilitate autentică sau de experiență conștientă ontologic coerentă. Evident, nu sunt de acord cu această linie retorică atunci când este aplicată procariotelor sau eucariotelor unicelulare, dar s-ar putea să aibă dreptate când vine vorba de regnul vegetal. Bănuiesc că în viitor vom descoperi că aceste diverse abilități florale pe care Baluska, Mancuso, Trewavas, Gagliano și alții le-au raportat sunt doar atât, acțiuni bio-mecanice necugetate, lipsite de subiectivitate. De asemenea, se va considera că plantele s-au angajat în ceea ce Janvier a numit un act de „degenerescență spectaculoasă” atunci când au abandonat conștiința, sensibilitatea pe care le-au păstrat ceilalți veri eucarioți din regnul animal.

Dacă am greșit, am greșit. Problema va fi rezolvată prin metode și cercetări științifice adecvate. Din perspectiva **CBC**, chiar nu contează dacă plantele au ceva asemănător experienței fenomenale. Cadrul general este valabil dacă există sau nu ceva asemănător cu a fi un cactus. Dar vor apărea alte probleme, cele care se referă la considerații etice. Mai simplu spus, vegetarienii și veganii care și-au adoptat modul de alimentație din motive etice se vor confrunta cu probleme.

Considerații etice

Modelul **CBC**, deoarece investește speciile de la procariote încoace cu stări subiective și o formă sau alta de sensibilitate, invită sau, poate mai bine, insistă asupra unei conversații privind etica, drepturile animalelor și (poate) drepturile plantelor. Dacă vom ajunge să demonstrăm în mod concludent că cadrul **CBC** este meritoriu și că toate speciile de animale, de la procariotele unicelulare primordiale încoace, sunt sensibile, au minți ontologic reale și experiențe subiective, nu putem scăpa cel puțin de o reflecție asupra preocupărilor etice care însoțesc o astfel de concluzie.

Dacă mă înșel și unele specii se dovedesc a fi simple sisteme bio-mecanice care își desfășoară repertoriile comportamentale fără nici măcar un gest de îngrijorare sau de încântare, atunci nu trebuie să ne facem griji în privința lor. În opinia economistului Yew-Kwan Ng, dacă nu au sensibilitate, nu au conștiință și, în lipsa acesteia, nu trebuie să fim preocupați de bunăstarea lor. Evident, aceasta nu este o problemă simplă. Nu am soluții, dar am gânduri, așa că permiteți-mi să vi le împărtășesc.

Primul punct pe care trebuie să îl abordăm este suferința. Am ridicat această problemă inițial în discuția noastră despre afirmația lui Brian Key (și a altora, precum Stuart Derbyshire) conform căreia peștii nu simt durere. Key și cei care apără această poziție își bazează concluzia pe argumentul că adevărata durere fenomenală depinde de o specie care împărtășește anumiți centri cerebrali și căi neuronale cunoscute ca fiind asociate cu experiența durerii la primate și mamifere. Conform acestei poziții, orice membru al unei specii căruia îi lipsesc acești centri și căi neuronale poate părea că suferă, dar aceste acțiuni sunt, din nou, simple reacții mecanice fără nicio suferință asociată. Un pește care plutește pe fundul unei bărci, din această perspectivă, nu suferă - reacționează doar într-un mod mecanic, lipsit de sensibilitate, la lipsa bruscă de oxigen. Poziția lui Key, după cum mulți au remarcat, este un vestigiu al argumentelor carteziene anterioare și, la fel ca analiza lui Descartes, acum aproape universal respinsă, invită și/sau justifică acțiuni care sporesc suferința.

Dacă abordarea **CBC** este corectă, dacă toate animalele, până la cele mai puțin complexe, au experiențe asemănătoare și de-a lungul unui continuum evolutiv cu ale noastre, dacă durerea lor se încadrează în aceeași categorie ontologică subiectivă cu durerea noastră, ce considerații suntem obligați să le acordăm? Ce modificări în comportamentul nostru față de toate speciile trebuie încurajate? Eticieni precum filosoful și militantul pentru drepturile animalelor Peter Singer susțin că avem obligația morală de a depune eforturi pentru a diminua suferința animalelor. Majoritatea eforturilor lui Singer vizează animalele „mari”, animalele de companie, cele din lanțul alimentar, cum ar fi porcii, vitele și oile, cele aflate în captivitate în circuri și grădini zoologice, cele a căror viabilitate ca specie este în pericol și cele utilizate în cercetarea științifică. Se manifestă puțină îngrijorare pentru insecte, cefalopode, nu prea multă pentru reptile și niciuna pentru cele unicelulare. De asemenea, se observă adesea că intrăm în contact cu aceste din urmă specii doar în circumstanțe întâmplătoare, neobișnuite și rare, astfel încât este mai puțin probabil să le provocăm durere sau răni, în timp ce industria alimentară, pescuitul, agricultura și practicile de creștere a animalelor sunt extrem de frecvente și implică un număr mare de organisme. În consecință, acțiunea umană poate provoca, și adesea provoacă, suferință, iar factorii economici și financiari acționează în moduri care compromit eforturile de promovare a bunăstării animalelor.

O perspectivă conexă este cea a filosofului Jonathan Birch de la London School of Economics, care a lucrat la dezvoltarea unui cadru în care *principiul precauției* ar trebui să fie aplicat atunci când se iau astfel de decizii morale și etice. *Principiul precauției*, pe care l-am abordat mai devreme, susține, în esență, că este mai bine să greșim în a nu provoca durere sau suferință, în special atunci când nu suntem siguri dacă durerea sau suferința este, de fapt, un aspect al circumstanțelor. Principiul, atunci când este aplicat la cazuri precum poziția lui Key privind durerea la pești, spune în mod clar că, dacă există îndoieli în această privință sau posibilitatea ca acțiunile noastre să provoace durere unui pește, poziția etică este să nu acționăm în acest mod.

Dar, după cum remarcă Birch, deoarece *principiul precauției* este invocat într-o multitudine de situații diferite, acesta tinde să fie aplicat cu o pensulă destul de largă. În opinia sa, este nevoie de un set de criterii care să poată fi utilizate pentru a identifica circumstanțele în care ar trebui să fie aplicat și cele în care nu trebuie. Pragul de aplicare pe care îl recomandă este sensibilitatea. În abordarea lui Birch, dacă se poate demonstra că un animal este sensibil, dacă „trece”, ca să spunem așa, testul Nagel, care arată că există ceva asemănător cu a fi din acea specie, atunci ar trebui considerat sensibil. Nu trebuie să știm cum este de fapt ca acel animal să simtă durere (și nu putem), ci doar să avem suficiente dovezi care să ne oblige să invocăm principiul. După cum a spus el:

Ar trebui să solicităm cel puțin un indicator credibil al sensibilității... Cel puțin, un indicator trebuie să fie un fenomen observabil pe care experimentele pot fi concepute pentru a-l detecta și trebuie să fie credibil faptul că prezența acestui indicator este explicată prin sensibilitate. (Subliniere adăugată.)

Birch recunoaşte că acest lucru aruncă mingea înapoi în terenul biologului şi/sau al psihologului, dar susţine că aceasta este singura abordare viabilă atunci când se iau în considerare legile, legislaţia şi reglementările care reglementează modul în care diverse specii sunt tratate în cadrul comerţului, sportului şi cercetării. În opinia sa, există suficiente dovezi din studiile empirice pentru a fi destul de siguri că crustaceele, cefalopodele şi peştii osoşi trec testul sensibilităţii. Interesant este faptul că Birch ia în considerare posibilitatea ca toate speciile de animale să fie sensibile, dar alege să nu meargă pe această cale - nu pentru că ar exista probleme etice dificile, ci pentru că principala sa preocupare este ce fel de legislaţie, ce forme de reglementare ar trebui să existe pentru a aborda problemele etice şi practice ale bunăstării animalelor. Cât de departe trebuie să mergem în domeniul simplităţii biologice înainte de a înceta să ne mai îngrijorăm cu privire la echilibrul dintre a nu face rău acelor specii şi a nu crea dificultăţi pentru noi şi pentru modul nostru de viaţă? Nu există un răspuns clar aici.

Cele mai primitive specii luate în considerare de Birch sunt nematodele şi muştele de fructe. În cazul în care, îşi face griji, ar trebui să stabilim că acestea sunt sensibile şi că intră sub incidenţa legislaţiei privind bunăstarea animalelor, acest lucru ar avea un impact devastator asupra cercetării biomedicale. În cele din urmă, Birch adoptă o abordare pragmatică a problemei, observând că există probleme etice şi probleme practice; există probleme filosofice şi probleme care ar trebui să guverneze reglementările şi acţiunile guvernamentale. Dar, deoarece aceste specii unicelulare sunt în centrul argumentelor aduse în apărarea abordării **CBC** şi se presupune că sunt sensibile, ne aflăm în aceeaşi situaţie în care am fost cu plantele, problemele trebuie, cel puţin, să fie discutate.

Folosind principiul pragmatic al precauţiei al lui Birch şi/sau ceea ce Singer numeşte o „etică practică”, punctul de plecare evident este de a face distincţia între acele procariote care ne sunt benefice şi cele care nu ne sunt deloc benefice. Unele bacterii, aşa-numitele specii probiotice, precum genurile *Lactobacillus* şi *Bifidobacterium*, oferă beneficii semnificative. Altele sunt neutre şi nu au un impact deosebit asupra vieţii umane. Altele sunt clar maligne, cum ar fi *Staphylococcus* şi Mycobacterium *tuberculosis*, care provoacă boli grave şi care, în ultimii ani, au evoluat cu tulpini rezistente la antibiotice, sporind dramatic ameninţarea la adresa vieţii mamiferelor. În cele din urmă, există specii precum *E. coli*, omniprezentă, ale cărei tulpini provoacă afecţiuni gastrointestinale grave, însă altele pot avea efecte benefice asupra tubului digestiv.

În mod clar, din perspectiva noastră centrată pe specie, ar trebui să ne dorim să aducem numai satisfacţie în viaţa oricărui *Lactobacillus* pe care l-am putea întâlni, dar utilizarea întregii game de arme antibiotice împotriva *Mycobacterium tuberculosis* şi cantitatea de suferinţă pe care am putea-o provoca membrilor acestei specii nu este, cu siguranţă, un factor atenuant - având în vedere reducerea suferinţei umane pe care ar aduce-o dispariţia lor. În casa noastră, furnicile, păianjenii, urechelniţele şi alte creaturi relativ inofensive, precum viespile, sunt uşor transportate afară. Nu este neobişnuit ca unul dintre noi să prindă delicat o urechelniţă în întredeschidere şi să o ducă în grădina de unde a venit. Cu toate acestea, ţânţarii sunt loviţi, iar melcii de grădină se îneacă în butoaie de bere (ceea ce, presupun eu, le diminuează suferinţa).

Spre deosebire de eticieni precum Singer, care este în principal vegan din motive etice, şi de cercetătorul cognitiv Stevan Hamad, care este un apărător pasionat al drepturilor animalelor şi este vegan de zeci de ani, eu sunt încă omnivor. Este important ca suferinţa să fie redusă; în casa noastră se consumă în principal alimente crescute organic şi crescute în aer liber şi susţin eforturile de reglementare pentru a face creşterea animalelor mai umană. Recunosc că există mai multe motive pentru care cineva ar putea deveni vegetarian. Un bun prieten evită carnea de pe uscat, dar consumă cu plăcere tot ce provine din ape. Altul este vegetarian, cu excepţia *scoicilor.* Amândoi spun că este o chestiune de pipăit în gură şi de gust. Alţii adoptă o dietă vegetariană sau vegană din motive de sănătate, considerând că eliminarea grăsimilor animale favorizează sănătatea şi o viaţă mai lungă şi

mai plăcută. În toate aceste cazuri, considerentele etice fie nu fac parte din arborele decizional, fie joacă un rol secundar. Alții, însă, adoptă o dietă vegetariană sau vegană din motive etice. Aceștia nu doresc să contribuie la suferință - despre care cred că este un rezultat direct al modului în care sunt crescute și sacrificate animalele destinate alimentației. Unii observă, de asemenea, că creșterea animalelor este un factor semnificativ în schimbările climatice. Ei doresc ca alimentele să fie produse în modul cel mai durabil și consideră că creșterea animalelor pentru hrană nu contribuie la atingerea acestui obiectiv de mediu. Expertul în energie Vaclav Smil remarcă faptul că, dacă toată lumea de pe planetă ar urma practica sa și ar limita consumul de carne la o singură masă pe săptămână, biosfera ar fi foarte bine. Mulți dintre cei care fac parte din mișcarea pentru protecția animalelor se referă la infractori ca la cei care exploatează animalele pentru cei trei F: Food, Fur, Fun (Mâncare, blană, distracție).

Dar există un adevăr brutal care trebuie recunoscut, chiar și de către cei care au ales un stil de viață vegetarian sau vegan din motive etice. Fiecare organism, în afară de prietenii noștri unicelulari care se hrănesc cu molecule inerte de zahăr, supraviețuiește ucigând și consumând alte organisme sau așteptând ca fapta să fie făcută de alții și apoi făcând curățenie. Așa se întâmplă când ai de-a face cu o planetă pe care viața se bazează pe carbon și apă. Atunci când aceste alte organisme sunt fie nebenefice, fie letale sau maligne pentru viața umană sau pentru alte forme de viață simțitoare, problema etică a provocării morții sau a durerii poate fi și de obicei este lăsată deoparte. Aceasta a fost/este abordarea standard în cadrul utilitarismului. Deși utilitarismul are (desigur) multe varietăți diferite, toate se concentrează pe stabilirea unui punct de echilibru în care acțiunile oamenilor sunt astfel încât utilitatea unui cadru, a unui mod de viață, a unui sistem social să crească. Pe scurt, atunci când se iau în considerare acțiunile față de animale care nu fac nimic care să ne diminueze experiențele și despre care se știe că sunt sensibile și capabile de suferință, ar trebui depuse eforturi care să sporească fericirea sau satisfacția generală și să diminueze sau să prevină suferința - în balanță cu costurile pe care astfel de acțiuni le pot avea asupra calității vieții oamenilor.

Revenind încă o dată la problema durerii la pești, din punctul de vedere pe care îl susțin aici, ar trebui să facem tot ce putem pentru a reduce suferința la pești sau la orice alt organism care poate simți durerea, deoarece, din perspectiva **CBC**, toată această durere se află pe același continuum evolutiv ca și a mea și a ta. O poziție utilitaristă pragmatică similară este recomandată atunci când luăm în considerare speciile utilizate în cercetare.

Dar plantele? Cum intră ele în acest cadru? Ca să fiu sincer, unul dintre motivele pentru care am susținut inițial că plantele nu sunt sensibile, că capacitatea fenomenală a strămoșului lor singular, eucariotul care a încorporat un procariot fotosintetizant, a adoptat viața unui tip nesensibil, fără sentimente și fără conștiință, a fost acela de a evita dilema de a trebui să le atrag atenția vegetarienilor etici că mănâncă alimente sensibile. Într-un mod compatibil cu argumentele lui Birch privind principiul precauției, plantele nu prezintă trăsăturile și comportamentele criteriale care ar determina un observator sensibil să le acorde sensibilitate. Prin urmare, nu trebuie să fim preocupați de bunăstarea lor.

Cu toate acestea, dacă ar fi prezentate suficiente dovezi care să susțină concluzia că plantele sunt conștiente, vor fi ridicate probleme etice legitime. Acestea nu au o soluție evidentă, alta decât cea oferită de Baluska: vegetarienii și veganii și-ar putea limita dieta la speciile de plante care doresc să fie consumate, cum ar fi cele care folosesc canalul alimentar al speciilor animale pentru a se înmulți și pentru a crește aria geografică a urmașilor lor.

În treacăt fie spus, voi observa că până acum nu am văzut această problemă dezbătută sau măcar ridicată în cadrul grupului de botaniști care susțin că plantele sunt sensibile. În cartea sa *The Cabaret of Plants* (Cabaretul plantelor), Richard Mabey susține că există motive pentru a crede că plantele „merită să primească același statut etic ca animalele”, dar nu merge până la capăt. El observă că dovezile recente care arată că plantele învață, au amintiri stabile și rezolvă probleme

susțin concluzia că acestea sunt inteligente - în sensul că se comportă într-un mod care exprimă inteligență. Dar este reticent când vine vorba de sensibilitate, subiectivitate, conștiință. Într-un interviu recent, el a afirmat că plantele prezintă trăsături care par a fi inteligente, dar fac acest lucru fără a avea la bază vreo sensibilitate care să se refere la sine.

Ceea ce sugerează noua botanică este că plantele sunt sensibile și rezolvă probleme, dar ocolesc nevoia de conștiință de sine - plantele, prin mijloace pe care nu le înțelegem încă pe deplin, sunt capabile să se comporte ca ființe inteligente.

Acest argument, desigur, este cel inspirat de Dennett, pe care îl apăr și eu, și anume că comportamentele manifestate de plante, care par a fi manifestări ale sensibilității, sunt doar reacții mecanice codificate genetic, desfășurate automat și fără sensibilitate sau experiență subiectivă. Până când vor fi aduse mai multe dovezi cu privire la chestiunea sensibilității plantelor, va trebui să rămânem aici.

Un scurt comentariu despre altruism și cooperare

Aș dori să arunc o privire puțin mai lungă asupra unei probleme pe care am trecut-o cu vederea mai devreme, dovezile a ceea ce unii ar putea considera comportament altruist la nivel unicelular. Aceasta va fi o secțiune scurtă. Altruismul este un subiect complex și ar fi nepotrivit să încercăm să-l abordăm aici în profunzime. Dar constatările pe care le-am analizat mai devreme au avut ceva provocator, în special acele situații în care s-a descoperit că coloniile bacteriene se comportă în moduri de cooperare care, cu siguranță, par să se potrivească cu caracteristicile general acceptate ale altruismului.

Criteriul obișnuit pentru ca un act să fie considerat altruist este ca acesta să aducă beneficii altei persoane, având în același timp un efect negativ asupra actorului. Subiectul este unul bogat cercetat în științele sociale, biologia evolutivă și în abordările filosofice ale eticii și raționamentului moral. Ca și în cazul conștiinței, abordarea obișnuită începe (și adesea se încheie) cu ființele umane și, în acest context, au fost identificate o serie de forme distincte de altruism. De obicei, accentul se pune pe probleme precum: Care acte sunt raționale, care nu sunt? Ce circumstanțe favorizează comportamentul altruist și care îl fac mai puțin probabil? Care sunt relațiile dintre agenții implicați în interacțiune care favorizează sau diminuează probabilitatea unui act altruist? Rudele sunt mai susceptibile să evoce altruismul decât neamurile? Ce rol joacă reciprocitatea, garantată sau doar presupusă? Grupurile se comportă diferit față de indivizii singuri? Ne observă cineva?

Deși aceste considerații și altele similare se află în afara domeniului procariotic care ne preocupă, pot fi ridicate probleme care sunt surprinzător de strâns legate, cum ar fi: Unde, în evoluție, vedem pentru prima dată altruismul autentic, în care un organism (sau un grup) acționează într-un mod care îi diminuează bunăstarea sau îi scade viabilitatea în beneficiul altui organism (sau grup)? Are altruismul, chiar și cel mai elementar, un element de reciprocitate? Există un anumit nivel de cooperare între indivizi? Este implicat un factor de rudenie?

Unele cercetări privind comunicarea în bacterii, în special cele ale grupului lui Gürol Süel de la Universitatea California din San Diego, pe care le-am analizat în *capitolul 4*, oferă răspunsuri tentante, deși provizorii, la aceste întrebări. După cum s-a menționat acolo, într-o serie de studii, bacteriile dintr-o colonie au fost observate ajustându-și metabolismul și ratele de diviziune celulară ca răspuns la semnalele moleculare de la bacteriile din alte părți ale colectivului. Celulele de la periferia unei colonii au acces la un mediu bogat în nutrienți, dar sunt expuse pericolului prădătorilor, antibioticelor și toxinelor. Celulele din centru se bucură de protecția locației lor, dar au acces diminuat la hrană. Dacă nivelul de nutrienți din centrul colectivului scade prea mult, celulele de acolo secretă o moleculă de semnalizare care, în esență, cere celulelor de la margini să încetinească, să nu se mai reproducă și să lase o parte din materialul bogat în nutrienți să curgă spre

interior. Atunci când celulele de la periferie se acomodează, ele sunt, de fapt, angajate într-un act altruist. Ele se comportă într-un mod care sporește bunăstarea colegilor lor de colonie, dar se pun pe ele însele în pericol. Pericolele potențiale ale prădătorilor și ale moleculelor nocive încă se ascund în mediul înconjurător, iar prin scăderea ratei metabolice și a diviziunii celulare, ele își diminuează șansele de a-și transmite genele în generația următoare.

Süel și colegii săi au examinat, de asemenea, aceste tipuri de comunicare între coloniile de bacterii, nu doar în interiorul acestora. Acolo au descoperit că funcționează un mecanism similar. În cazul în care nivelul nutrienților din zona unei colonii scade, colectivul secretă ioni care se deplasează ca o undă prin spațiul dintre cele două, informând celălalt grup că „lucrurile nu merg bine pe aici". Cel de-al doilea grup răspunde cu ceea ce pare a fi același tip de act altruist, reducându-și rata de diviziune celulară și funcțiile metabolice. Atunci când oamenii lui Süel au ajustat distribuția nutrienților în mediu, astfel încât al doilea colectiv să se confrunte acum cu o penurie de hrană, ei au emis același semnal, iar prima colonie, acum refăcută, a făcut o adaptare similară. Controlând nivelurile de nutrienți în locuri diferite, ei au reușit să manipuleze semnalizarea și diviziunea celulară în ambele colective.

În prima serie de experimente cu colonii separate, cele două grupuri erau din aceeași specie de bacterii. Într-un studiu ulterior, cercetătorii au constatat că aceeași formă de altruism apare și atunci când cele două colective de bacterii provin din specii diferite. Luate împreună, aceste constatări constituie un argument solid pentru existența comportamentului altruist la procariote.

Cu toate acestea, descoperirea care vizează mai multe specii este complicată. Este posibil ca mecanismul care acționează aici să fie unul care funcționează independent de specia coloniilor care cooperează și, bineînțeles, nu știm dacă organismele unicelulare precum bacteriile se deranjează să noteze identitatea de specie a vecinilor lor. Atunci când te reproduci prin diviziune celulară, identificarea organismelor conspecifice nu este o problemă. Nu este clar că acest din urmă studiu dezvăluie ceva mai mult decât primul. Grupul lui Süel s-a referit la efectul general ca la un fel de „împărțire a timpului" și, din cauza schimburilor de roluri, constatările lor arată una dintre cheile frecvent citate ale unui altruism interdependent, reciprocitatea. „Voi face acest lucru pentru a te ajuta cu condiția să știu că vei face același lucru pentru mine atunci când voi avea nevoie de ajutorul tău".

Susțin eu că bacteriile se angajează într-un comportament ontologic solid și de înaltă moralitate? Nu. Dar observ că comportamentul altruist există în tot spectrul speciilor, că probleme precum aceasta ar trebui tratate la fel ca și conștiința - apărând pentru prima dată în funcțiile speciilor primitive și evoluând spre elementele multiple și complexe observate în alte organisme, pe măsură ce selecția naturală își face treaba. Atunci când ființele umane se pun pe ele însele în pericol pentru binele altora, ele aduc în acest act toate trăsăturile raționalității, bunătatea inimii, dorința de a-i ajuta pe alții, oportunitățile de a-și stimula ego-ul, elemente ale teoriei jocurilor, credința în principiul reciprocității, credința într-un Dumnezeu divin omniscient - lista factorilor care contribuie la probabilitatea ca cineva să se angajeze într-un act de altruism este lungă și variată. Niciunul dintre aceștia nu se regăsește la bacterii, dar acestea par să fie niște creaturi mici și altruiste și, deoarece aceste comportamente par să fie codificate în genele procariotelor, este posibil ca ele să fie bine fixate în ADN-ul multor alte specii. La fel ca în cazul conștiinței și al vieții mentale, este posibil să avem în vedere prima apariție a altruismului reciproc și rădăcinile biologice ale tuturor celorlalte forme. Există, de asemenea, un ușor imbold sociopolitic în urma acestor descoperiri pe care este greu să nu-l observăm: bacteriile nu par a fi fane ale lui Ayn Rand.

Așa cum am văzut în cazul cercetării în domeniul inteligenței artificiale și al neuroștiințelor, rapoartele care provin din aceste laboratoare de biologie nu sunt prezentate în mod obișnuit ca demonstrații de altruism reciproc. Termenul este rar utilizat în literatura de specialitate din domeniul biologiei celulare. După cum am observat mai devreme, oamenii de știință fac știință și o fac în mare măsură independent de implicațiile pe care descoperirile lor le-ar putea avea pentru

problemele care îi preocupă pe cei din departamentul de filosofie. Cred că acest lucru este în regulă. Dar distracția începe atunci când filosofii își dau seama de rezultatele lor.

Câștiguri secundare

Ei bine, am cam terminat. Așa cum am spus, aceasta a fost planificată să fie o carte scurtă. Cărțile scurte au o șansă mai mare de a fi citite. Am avut o teză de prezentat. Se bazează pe niște principii foarte simple, pe care le-am expus, sper, într-o manieră clară. Este, de asemenea, radicală prin faptul că se situează în afara curentului principal din biologia evoluționistă, neuroștiințele cognitive și filosofia minții. A fost nevoie de ceva efort pentru a oferi ceea ce cred eu că este o apărare decentă. Sunt sigur că voi ști destul de curând dacă am reușit să fac acest lucru.

Permiteți-mi să închei prin a sublinia faptul că modelul **CBC**, pe lângă prezentarea unui cadru biofizic nou, are o serie de ceea ce cred că putem numi câștiguri secundare, consecințe neintenționate de tip productiv. Unele dintre acestea sunt de natură filozofică, în sensul că oferă o perspectivă diferită asupra problemelor dezbătute în prezent în domeniu. Unele ar trebui să încurajeze biologii evolutivi să pună alte tipuri de întrebări decât cele considerate în mod obișnuit. Altele sunt de natură metodologică în sensul că invită la modalități diferite de abordare a problemelor vechi. În cea mai mare parte - trebuie să fiu sincer aici - acestea nu erau în gândirea mea când am început această călătorie. Metaforic vorbind, ele s-au pus pe masă în fața mea pe măsură ce proiectul se dezvolta.

- *Problema dificilă* este rezolvată - nu „soluționată”, ci „rezolvată”. Ea este încă acolo.[82] Va trebui în continuare să stabilim relațiile dintre funcțiile creierului și experiența umană, va trebui în continuare să găsim centrele neuronale și zonele cerebrale care fac parte integrantă din funcțiile cognitive și afective. Dar acest lucru poate fi realizat în cadrul a ceea ce Thomas Kuhn a numit în mod celebru „știința normală”. Problemele sunt identificate, se generează ipoteze, se fac experimente și se ajunge la concluzii. Tot nu va fi ușor, dar cel puțin nu vom mai fi agățați de petarda decalajului explicativ.

- Problemele legate de decalajul explicativ dispar. Din perspectiva **CBC**, nu există niciun motiv să ne imaginăm că există un decalaj explicativ care ar trebui să ne preocupe. După cum s-a menționat mai sus, o multitudine de astfel de lacune există în marea întindere a științei și au existat întotdeauna. Ori de câte ori ne confruntăm cu o situație în care nu avem o înțelegere solidă a legăturilor cauzale dintre factori, când nu ne este clar cum un X (sau un grup de X) cauzează un Y (sau variații ale lui Y), cum se produce un proces complex sau cum funcționează un mecanism complicat, avem de-a face cu o „lacună explicativă”. Acesta este modul în care funcționează știința, iar descoperirea relațiilor este modul în care funcționează cel mai bine. Lacuna epistemică care se profilează, ca o prăpastie, în gândirea multor filosofi nu mai este ceva care să ne preocupe teribil.

- Problema celorlalte minți este neutralizată. Deși trebuie să recunoaștem punctul de vedere al lui Nagel că suntem cu toții prinși în capcana perspectivei noastre la persoana întâi și că nu putem ști cu adevărat cum este mintea unei alte persoane (sau a unui alt animal), nu trebuie să ne facem griji în această privință. Trebuie doar să presupunem că fiecare specie este sensibilă, să lăsăm datele să ne spună ce tipuri de comportamente sunt posibile și să facem presupuneri educate cu privire la ce tipuri de experiențe subiective sunt prezente. Ne ocupăm de știință fără a ne lăsa prinși în capcana nerezolvabilului. După cum am menționat anterior, cel mai bun răspuns la problema celorlalte minți ar putea fi „și ce dacă”.

- Misterianismul este evitat. Eșecul de a rezolva *problema dificilă*, de a afla cum creierul creează mintea, i-a condus pe mulți în tărâmuri filosofice nefericite, în zone abstracte din punct de vedere conceptual, pe căi fără ieșire. Unii au îmbrățișat una sau alta dintre versiunile panpsihismului

82 Deși un recenzent al prospectului pentru această carte s-a întrebat dacă nu cumva abordarea **CBC** nu a „dizolvat-o”.

care conferă întregii realități una sau alta dintre formele evazive ale conștiinței. Alții au îmbrățișat oricare dintre cele câteva versiuni vag legate ale dualismului, în care mentalul era, cumva, plasat pe un plan ontologic distinct, separat de material și, probabil, urmând legi și principii separate. Alții pur și simplu au aruncat prosopul și au declarat că problema nu poate fi rezolvată, cel puțin de către noi - un act care mi se pare cel mai misterios dintre toate. Poziția **CBC** este clar disociată de acestea și de alte forme de gândire misterioasă.

- Centrismul speciei este diminuat și este încurajat un antropomorfism pragmatic, critic. Sunt cu siguranță de acord cu nevoia pe termen lung de a lucra la relațiile complexe dintre mecanismele cerebrale și actele cognitive și emoționale la oameni. Dar tendința neuroștiințelor și a filosofilor minții de a începe și de a termina prea des cu viața mentală umană nu a fost cea mai productivă dintre strategii. Acest egoism de specie a îngreunat aprecierea faptului că toate actele noastre mentale bine puse la punct au rădăcini evolutive în alte specii, că practic fiecare lucru interesant pe care îl facem își are începuturile în altă parte în panoplia filogenetică. Suntem speciali, dar nu chiar atât de speciali. Suntem doar un punct final de moment într-o călătorie evolutivă care va continua, cu sau fără noi.

- Un antropomorfism critic și măsurat, unul îmbrăcat în bun simț, este recunoscut pentru contribuțiile sale potențiale la aceste probleme mentale. Prin sublinierea comunității speciilor, căutarea rădăcinilor biologice și comportamentale comune ale formelor și funcțiilor umane poate fi efectuată mai eficient.

- O eroare de categorie este evitată. Acest mic bonus intelectual vine odată cu precedentul. Prin renunțarea la abordarea centrată pe *H. sapiens* evităm tendința nefericită de a considera conștiința umană ca având cumva o ontologie distinctă, o existență separată conceptual de experiențele mentale ale altor specii. În cadrul abordării **CBC**, conștiința umană nu este un tip distinct de conștiință. Ea este un simbol și se află la polul îndepărtat al bogăției și complexității, dar împărtășește tipul său de bază, fundamental, cu toate celelalte forme de sensibilitate.

- O poziție fizicistă, materialistă cu privire la viața mentală este salvată de atacul asupra ei din mai multe părți. Problemele cu care ne confruntăm, filosofice și empirice, sunt reale, dar mult mai ușor de rezolvat decât cele cu care ne-am confruntat atunci când principala întrebare pusă a fost: „Cum poate simpla materie să creeze mentalul?”

- Se pune accentul pe un reducționism pragmatic. Parcurgerea traseului **CBC** va încuraja căutarea elementelor de bază care stau la baza conștiinței, a sensibilității. Este un reducționism care se concentrează asupra întrebărilor puse. În cadrul **CBC**, întrebarea principală este „Cum apare sensibilitatea la speciile unicelulare?” Dacă, într-un viitor oarecare, vom cunoaște răspunsul, ar putea fi util să coborâm un nivel sau două și să vedem ce roluri joacă elementele mai de bază. Aceasta va depinde de întrebările la care trebuie să răspundem.

- William de Occam, Conwy Lloyd Morgan și toți cei care apreciază valoarea (din nou, pragmatică) a parcimoniei sunt susținuți. Dacă poziția **CBC** rezistă unei examinări empirice, răspunsul la marile întrebări privind originile vieții mentale va fi exprimat în termenii cei mai simpli posibili. Desigur, problemele vor exista în continuare, dar modul în care le putem conceptualiza și tipurile de strategii de cercetare pe care le vom adopta vor fi diferite - și mai simple.

- O nouă perspectivă asupra unei varietăți de probleme auxiliare care implică altruismul, etica și bunăstarea animalelor este pusă pe masă alături de numeroasele perspective examinate în prezent. După cum s-a menționat, încurajarea cercetătorilor să includă puncte de vedere noi și uneori radicale încurajează o creativitate emergentă. Ideile și perspectivele care se ascund în prezent sub pragul de recunoaștere sunt pregătite.

Rezumat

Deci am terminat. Nu știu dacă am convins pe cineva. O parte din mine speră că da, deoarece cred că modelul *Bazei Celulare a Conștiinței* invită la o nouă modalitate de abordare a

multor probleme filosofice, precum și a celor paralele din neuroștiințele cognitive și biologia celulară. Dar o altă parte din mine speră că nu am convins total mai mult de câțiva. În domenii atât de dificile precum cele pe care le-am abordat, se fac progrese, nu atunci când există un acord deplin, ci atunci când există dezbateri, dezacorduri, dispute sănătoase. Întotdeauna este mai util să avem oameni de știință și filozofi dedicați care să se confrunte, mai degrabă decât să se simtă confortabil. Întotdeauna mi-a plăcut o luptă intelectuală bună, un duel iute cu cei care nu au fost de acord cu mine. Nu mai fac cercetare (am închis laboratorul în 2005), dar încă mai scriu, încerc să rămân implicat și aproape toate lucrările mele recente provin dintr-o dispută sau alta cu un gânditor apreciat și respectat.

În cele din urmă, este clar că dacă există vreun câștig epistemic în adoptarea poziției **CBC** este, în analiza finală, o problemă empirică, nu una filosofică. Cineva va trebui să analizeze foarte atent genetica, microbiologia și nișele ecologice ale organismelor unicelulare pentru a stabili dacă acestea au rădăcinile biomoleculare, genetice ale sensibilității. Dacă da, bine. Dacă nu, bine, și asta este bine. Așa cum am afirmat la început, am început acest proiect pregătit să se demonstreze că mă înșel.

Cu siguranță, nu va fi o surpriză faptul că eu cred că am descoperit ceva important aici.

Anexa A

Un exercițiu de lexicografie: Definirea(?) conștiinței

Așa cum am menționat în Prefață și am presărat pe parcursul cărții, au existat motive pentru a alege calea nedefinitivă atunci când am avut de-a face cu termenul nostru central, **conștiința**. Unul dintre acestea a fost ceea ce am găsit în timp ce făceam diligența necesară, căutând o definiție utilă a conceptului nostru central. M-am gândit că un loc rezonabil pentru a începe acest exercițiu este să caut definiții „oficiale" ale termenului în literatura de referință evaluată de colegi. Cel mai „oficial" loc pe care îl cunosc și care se potrivește acestei descrieri este *Penguin Books Dictionary of Psychology* și, din moment ce eu l-am scris, m-am simțit confortabil să încep de aici - chiar dacă, trebuie să recunosc, atunci când am decis să arunc o privire, nu-mi puteam aminti cu adevărat ce scrisesem. M-am dus inițial la prima ediție, care a fost publicată la mijlocul anilor 1980, atunci când psihologia cognitivă ieșea din zonele periferice și devenea rapid abordarea dominantă în psihologie. Conștiința era dată acolo ca:

> *1. În general, o stare de conștiință; o stare de a fi conștient. Aceasta este cea mai generală utilizare a termenului și este prevăzută în fraze precum „și-a pierdut cunoștința".*
> *2. Un domeniu al minții care conține senzațiile, percepțiile și amintirile de care cineva este momentan conștient; adică acele aspecte ale vieții mentale prezente la care cineva este atent (vezi atenție).*
> *3. Acea componentă a minții disponibilă pentru introspecție. Acest sens se găsește în scrierile mai vechi ale structuraliștilor și ale altor introspecționiști.*
> *4. În psihanaliză, conștientul.*

Termenul are o istorie plină de conflicte. Acesta a fost reprezentat uneori ca punctul central al psihologiei (de exemplu, structuralismul), iar alteori a fost interzis din lexiconul psihologului ca reprezentând nimic mai mult decât resturile epifenomenale ale activității corporale (de exemplu, behaviorismul). Cu toate acestea, fascinația continuă față de aceasta provine din sentimentul convingător că conștiința este una dintre trăsăturile definitorii ale speciei noastre: că a fi om înseamnă a poseda nu numai conștiința de sine, ci și capacitatea și mai remarcabilă de a scana și de a revizui mental ceea ce conștientizăm. (Notă: mențiunea de aici nu este pentru a sublinia. Ea semnalează pur și simplu că termenul este definit în altă parte în dicționar).

Intrarea a continuat cu discutarea altor aspecte conotative, dar aceasta ne oferă o idee destul de bună despre modul în care termenul a fost utilizat în psihologie. Și-a făcut treaba și a abordat majoritatea problemelor, dar nu a făcut dreptate în ceea ce privește gama de probleme aflate în discuție în prezent. De asemenea, intrarea a fost complet centrată pe specie. Se referea doar la conștiința umană și trata termenul ca și cum ar denota funcții și procese întâlnite doar la *Homo sapiens*. Reflectam, desigur, prejudecățile din domeniu. Douăzeci de ani, trei ediții și doi noi coautori mai târziu, cea de-a patra ediție (acum cu Emily Reber și Rhianon Allen) prezintă în continuare aceleași teme. Există mai mult material actualizat, dar definiția reflectă încă o poziție centrată pe specii. Dacă va exista vreodată o a cincea ediție, această viziune îngustă va fi schimbată.[83]

Cum rămâne cu principalul nostru concurent, *Macmillan Dictionary of Psychology?* Cuvântul despre conștiință a fost scris de regretatul meu prieten și coleg lexicograf Stuart Sutherland. Aici există la fel de puțină asistență. Sutherland a scris:

> *Faptul de a avea percepții, gânduri și sentimente; conștiință. Termenul este imposibil de definit decât în termeni care sunt de neînțeles fără o înțelegere a ceea ce înseamnă conștiința. Mulți cad în capcana de a confunda conștiința cu conștiința de sine - pentru a fi conștient este necesar doar să fii conștient de lumea exterioară. Conștiința este un fenomen fascinant, dar evaziv; este imposibil de precizat ce este, ce face sau de ce a evoluat. Nu s-a scris nimic demn de citit despre ea.*

83 Penguin ne-a anunțat că este extrem de puțin probabil să existe ediții viitoare. Motivul? Evident. Internetul.

Abordarea lui Sutherland a făcut mai degrabă să pară că conştiinţa este ca o răceală. Nu poţi înţelege cu adevărat cum este sau cum te simţi dacă nu ai avut una. Inteligent, dar nu foarte lămuritor - deşi înclinaţia lui Sutherland către aspectul subiectiv, experienţial era tipică pentru modul în care majoritatea foloseau termenul. Această ultimă replică a fost, pentru cei care l-au cunoscut, clasică pentru Sutherland. De asemenea, a făcut ca definiţia sa să fie cu uşurinţă cea mai frecvent citată.

Aceste două eforturi au venit din partea psihologilor cu o orientare cognitivă. Este posibil ca eforturile psihologilor să se dovedească insuficiente, având în vedere că avem tendinţa de a avea propriile noastre opinii asupra lucrurilor mentale. Aşadar, să ne întoarcem la enciclopedia de opt volume a filosofiei, compilată şi editată de un alt vechi prieten de mult dispărut, fostul meu coleg de la Brooklyn College, filosoful Paul Edwards. Paul s-a adresat filozofului Charles Landesman pentru a se ocupa de intrarea referitoare la conştiinţă, iar Landesman a furnizat imediat mai multă căldură decât lumină:

Termenul „conştiinţă" apare în filozofie, psihologie şi vorbirea curentă cu o varietate de sensuri diferite. În acest articol vor fi discutate acele semnificaţii care sunt relevante pentru formularea mai multor probleme filosofice recurente.

Din nefericire, aceste „semnificaţii" s-au dovedit a fi categorii de referinţă care i-au permis lui Landesman să se sustragă de la toate exigenţele definitorii. El a început prin a trata conştiinţa ca şi cum ar putea (ar trebui?) să fie considerată drept „cunoaştere de sine" şi apoi a oferit o imagine de ansamblu sensibilă din punct de vedere istoric a acestei probleme, cu discuţiile aferente despre introspecţie, stări mentale şi dacă aceste stări de conştiinţă sunt sau nu alcătuite din „chestii" separate, nemateriale (adică problema dificilă înainte ca aceasta să fie denumită astfel). Nu că intrarea lui Landesman nu ar fi fost academică. A fost. Cu toate acestea, fără numeroasele conexiuni, ceea ce a rămas nu a fost mult mai mult decât conotaţiile psihologiei populare comune în domeniu şi, cu siguranţă, nu ne-a oferit o definiţie.

Următorul loc unde am verificat a fost intrarea în respectatul *Oxford Dictionary of Philosophy* al lui Simon Blackburn. Din păcate, era foarte asemănătoare cu cea a lui Landesman. Când i s-a aruncat o lovitură dură în sus şi înăuntru şi Blackburn, de asemenea, s-a ferit, a fugit direct din careul de bătaie. Definiţia lui nu era deloc o definiţie. Intrarea a început cu:

Posibil cea mai provocatoare şi omniprezentă sursă de probleme din întreaga filosofie.

Apoi a descris şi a conturat aceste probleme fără a face nicio încercare de a defini termenul în sine.

Dar laşitatea lexicografică de acest fel nu este nimic în comparaţie cu *Penguin Dictionary of Philosophy* editat de Thomas Mautner şi aflat acum la a doua ediţie. Ai putea crede că tu sau cineva a făcut o eroare de sortare alfabetică, dar, oricât ai căuta, nu vei găsi nicio intrare pentru conştiinţă... Şi nimic pentru inconştient, sau experienţă, sau conştiinţă. Fascinant cum un termen considerat de Blackburn drept „cea mai provocatoare şi omniprezentă sursă de probleme din întreaga filosofie" poate fi complet neglijat de Mautner.

Începând din 1995, Universitatea Stanford a început publicarea, editarea şi actualizarea periodică a *Enciclopediei Stanford de Filosofie*, accesibilă gratuit. Aceasta a devenit una dintre cele mai respectate surse online de informaţii generale privind conceptele şi termenii din domeniu (şi unul dintre motivele pentru care dicţionarele pe suport de hârtie nu prea mai zboară de pe rafturile librăriilor). Cuvântul despre conştiinţă începe într-un mod la care ar trebui să fie pregătiţi cei care sunt încă cu mine aici:

Poate că niciun aspect al minții nu este mai familiar sau mai derutant decât conștiința și experiența noastră conștientă de sine și de lume. Problema conștiinței este, fără îndoială, problema centrală a teoriei actuale despre minte. În ciuda lipsei unei teorii a conștiinței, există un consens larg, chiar dacă nu universal, conform căruia o descriere adecvată a minții necesită o înțelegere clară a acesteia și a locului său în natură. Trebuie să înțelegem atât ce este conștiința, cât și cum se raportează aceasta la alte aspecte ale realității, care nu sunt conștiente.

Ceea ce urmează urmărește aceeași cale familiară, numai că, fiind vorba de *Stanford Encyclopedia of Philosophy*, este mult mai extinsă decât majoritatea surselor. De fapt, restul intrării are peste 19 000 de cuvinte și explorează o mulțime de probleme specifice, teorii, dispute și rezolvări ocazionale în domeniu, fără a sugera, desigur, vreun efort de definire.

Pagina Wikipedia spune o poveste similară, deși are o frază revelatoare în intrarea sa destul de extinsă:

În ciuda dificultăților de definire, mulți filosofi cred că există <u>o intuiție</u> de bază larg împărtășită cu privire la ce este conștiința. (Subliniere adăugată).

Împărtășesc acest sentiment și, desigur, acesta este motivul principal pentru care am adoptat poziția psihologiei populare în această carte.

Căutările în scrierile cercetătorilor individuali care fac cercetări asupra conștiinței au dus cam la aceeași lipsă de claritate a sensului. Nici cartea respectată *Consciousness Reconsidered* a lui Owen Flanagan, nici controversata *The Conscious Mind* a lui David Chalmers nu încearcă în mod serios o definiție. În *The Nature of Consciousness: Philosophical Debates*, editat de Ned Block, Flanagan și Güven Güzeldere, niciunul dintre numeroșii autori nu este deosebit de util; la fel și cei care au contribuit la volumul interdisciplinar editat de Jonathan Cohen și Jonathan Schooler, *Scientific Approaches to Consciousness*.

Devenise amuzant. Indiferent cât de departe mergeam, nu găseam nicăieri un ajutor lexicografic autentic. Nu era ca și cum toți autorii evitau problema în mod conștient (!) - deși era destul de clar că nu puțini o făceau. Și nici nu era vorba de faptul că definițiile care erau furnizate nu erau satisfăcătoare. În puținele cazuri în care cineva a încercat cu adevărat să se confrunte cu semnificația termenului, acesta a produs de obicei: (a) variații ale „definițiilor” oferite de Sutherland și Blackburn, adică liste de probleme care au apărut atunci când s-a luat în considerare conștiința, sau (b) plângeri cu privire la faptul că nimeni nu are cu adevărat o idee coerentă despre ce este conștiința, în afară de acea idee personală pe care fiecare dintre noi o are cu privire la ceea ce naiba este cu siguranță conștiința fiecăruia dintre noi.

John Searle, deloc surprinzător, a abordat, de asemenea, problema definiției și, deloc surprinzător, consideră că are o modalitate de a o rezolva. Din cartea sa, *The Mystery of Consciousness*:

Se spune adesea că „conștiința” este îngrozitor de greu de definit. Dar dacă vorbim despre o definiție în termeni de bun simț, suficientă pentru a identifica ținta investigației, spre deosebire de o definiție științifică precisă de tipul celei care apare de obicei la sfârșitul unei investigații științifice, atunci cuvântul nu mi se pare greu de definit. Iată definiția: Conștiința constă în stări și procese interioare, calitative, subiective, de sensibilitate sau conștientizare. Conștiința, astfel definită, începe atunci când ne trezim dimineața dintr-un somn fără vise - și continuă până când adormim din nou, murim, intrăm în comă sau devenim în alt mod „inconștienți”. Ea include toată varietatea enormă a conștiinței pe care o considerăm caracteristică vieții noastre în stare de veghe. Acesta include orice, de la simțirea unei dureri, la perceperea vizuală a obiectelor, la stări de anxietate și depresie, la rezolvarea de puzzle-uri cu cuvinte încrucișate, la jocul de șah, la

încercarea de a vă aminti numărul de telefon al mătușii dumneavoastră, la discuții despre politică sau la dorința de a fi în altă parte. Conform acestei definiții, visele sunt o formă de conștiință, deși, desigur, sunt în multe privințe foarte diferite de conștiința trează.

Acest lucru nu este rău, deloc. Este un Searle de epocă, cu accentul său pe demonstrabilul empiric, dar atunci când despachetăm acest paragraf, se dovedește a fi de fapt un compendiu al acelor evenimente, stări și experiențe mentale care, în opinia sa, ar trebui să fie incluse sub termenul umbrelă în sine - care este cam același loc în care am ajuns atunci când am îmbrățișat sensul psihologiei populare. Searle, desigur, recunoaște acest lucru atunci când aruncă această replică: "... *vorbim despre o definiție în termeni de simț comun, suficientă pentru a identifica ținta investigației*".

Desigur, nimic din această agitație denotativă nu este nou. George Miller, unul dintre primii care au organizat un program serios de cercetare în domeniul psihologiei cognitive, s-a frământat cu privire la modul în care acest subiect preferat al său a fost abuzat. El și-a manifestat nefericirea în modul său obișnuit, elegant. Conștiința, pentru el, devenise:

Un cuvânt uzat neted de un milion de limbi. În funcție de figura de stil aleasă, este o stare de a fi, o substanță, un proces, un loc, un epifenomen, un aspect emergent al materiei sau singura realitate adevărată.... Poate că devenim confuzi pentru că, ori de câte ori ne gândim la conștiință, suntem înconjurați de ea și nu putem decât să ne imaginăm ce nu este conștiința.

Este demn de remarcat faptul că Miller a scris aceste rânduri pe vremea când psihologia abia începea să iasă din deceniile de dominație a behaviorismului - și, în mod amuzant, le-a scris în timp ce el și nepocăitul B. F. Skinner, care considera astfel de subiecte ca fiind inadecvate pentru atenția oamenilor de știință serioși, erau colegi în același departament la Harvard.

Această încurcătură denotativă și conotativă contorsionată îmi aduce, sper, puțină simpatie pentru faptul că am renunțat la orice eforturi serioase de a „la naiba, omule, definește-ți termenii" sau de a urmări semnificațiile multor altora, cum ar fi *minte, subiectiv, fenomenal, sensibilitate, conștiință sau sine și conștiință de sine* - toate acestea au fost, de secole, studiate de mulți.

Privind retrospectiv acest material, se poate observa tendința puternică de a lega diferitele semnificații de experiența umană, de tipurile noastre de subiectivitate, de tipurile noastre de minți. În esență, toți și-au asumat sarcina de a defini conștiința pornind de la prezumția că aceasta este o stare mentală experimentată în primul rând de membrii speciei noastre. Deși este adevărat că au existat câteva eforturi, cum ar fi cele ale lui Donald Griffin, pe care le-am discutat în *capitolul 2*, care au abordat posibilitatea minților animale, a fost rar să găsim un efort de definire care să fie lipsit de narcisismul speciei. Și chiar și atunci când posibilitatea conștiinței la alte specii a fost considerată plauzibilă, starea subiectivă a fost încă privită în contextul sau prin comparație cu experiența umană. Deoarece modelul **Bazei Celulare a Conștiinței** se concentrează asupra apariției inițiale a sensibilității, atunci când totul a fost spus și făcut, părea cel mai înțelept să se recurgă la sensurile psihologiei populare ale numeroșilor termeni utilizați ca sinonime libere pentru conștiință. Ceea ce, desigur, este ceea ce au făcut majoritatea celor care lucrează în numeroasele domenii care se suprapun și se întrepătrund și care se preocupă de acest subiect.

Anexa B

Comentariu la Declarația de la Cambridge privind conștiința

După cum s-a menționat în prefață, Declarația de la Cambridge a fost însoțită de o serie de comentarii prezentate de participanții la conferință. Acestea oferă o bună imagine de ansamblu a tipurilor de probleme ridicate atunci când se caută neurocorelații ai conștiinței la oameni și se descoperă dovezi ale acesteia la alte specii și sunt reproduse aici.

- Domeniul cercetării conștiinței evoluează rapid. Au fost dezvoltate numeroase tehnici și strategii noi pentru cercetarea pe animale umane și non-umane. În consecință, din ce în ce mai multe date devin disponibile, ceea ce necesită o reevaluare periodică a preconcepțiilor anterioare în acest domeniu. Studiile pe animale non-umane au arătat că circuitele cerebrale omologe corelate cu experiența și percepția conștientă pot fi facilitate și perturbate selectiv pentru a evalua dacă sunt de fapt necesare pentru aceste experiențe. În plus, la om, noi tehnici neinvazive sunt ușor disponibile pentru a studia corelațiile conștiinței.

- Substraturile neuronale ale emoțiilor nu par a fi limitate la structurile corticale. De fapt, rețelele neuronale subcorticale activate în timpul stărilor afective la om sunt, de asemenea, extrem de importante pentru generarea comportamentelor emoționale la animale. Trezirea artificială a acelorași regiuni ale creierului generează comportamente și stări sufletești corespunzătoare atât la oameni, cât și la animalele non-umane. Oriunde în creier se evocă comportamente emoționale instinctuale la animalele non-umane, multe dintre comportamentele care rezultă sunt în concordanță cu stările de simțire experimentate, inclusiv acele stări interne care sunt recompensatoare și punitive. Stimularea cerebrală profundă a acestor sisteme la om poate genera, de asemenea, stări afective similare. Sistemele asociate cu afecțiunea sunt concentrate în regiuni subcorticale în care abundă omologiile neuronale. Animalele tinere umane și non-umane fără neocortex păstrează aceste funcții creier-minte. În plus, circuitele neuronale care susțin stările comportamentale / electrofiziologice de atenție, somn și luare a deciziilor par să fi apărut în evoluție încă de la radiația nevertebratelor, fiind evidente la insecte și moluște cefalopode (de exemplu, caracatița).

- Păsările par să ofere, prin comportamentul, neurofiziologia și neuroanatomia lor, un caz izbitor de evoluție paralelă a conștiinței. Dovezi ale unor niveluri de conștiință apropiate de cele umane au fost observate cel mai dramatic la papagalii gri africani. Rețelele emoționale și microcircuitele cognitive ale mamiferelor și păsărilor par să fie mult mai omologe decât se credea anterior. În plus, s-a constatat că anumite specii de păsări prezintă modele neuronale de somn similare cu cele ale mamiferelor, inclusiv somnul REM și, așa cum s-a demonstrat la cintezele zebră, modele neurofiziologice despre care se credea anterior că necesită un neocortex de mamifer. În special, în studiile de autorecunoaștere în oglindă, s-a demonstrat că vrăbiile prezintă asemănări izbitoare cu oamenii, maimuțele mari, delfinii și elefanții.

- La om, efectul anumitor halucinogene pare a fi asociat cu o perturbare a procesării corticale de *feed forward* și *feedback*. Intervențiile farmacologice la animalele neumane cu compuși despre care se știe că afectează comportamentul conștient la om pot duce la perturbări similare ale comportamentului la animalele neumane. La om, există dovezi care sugerează că conștientizarea este corelată cu activitatea corticală, ceea ce nu exclude posibile contribuții din partea procesării subcorticale sau corticale timpurii, ca în cazul conștientizării vizuale. Dovezile conform cărora sentimentele emoționale ale oamenilor și ale animalelor non-umane provin din rețele cerebrale subcorticale omologe oferă dovezi convingătoare pentru *qualia* afectivă primară comună din punct de vedere evolutiv.

TRADUCERE REALIZATĂ CU:

DeepL Translator